

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:16 ; Search time 4521.53 Seconds
(without alignments)
16.136 Million cell updates/sec

Title: US-09-164-714-13

Perfect score: 18
Sequence: 1 ggtgtgtgtgtgtgtt 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *
44: gb_est25: *

45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_est48: *
83: gb_est49: *
84: gb_est50: *
85: gb_est51: *
86: em_est34: *
87: em_est35: *
88: em_est36: *
89: em_est37: *
90: gb_est52: *
91: gb_est53: *
92: gb_est54: *
93: gb_est55: *
94: gb_est56: *
95: em_est38: *
96: em_est39: *
97: em_est40: *
98: em_est41: *
99: em_est42: *
100: em_est43: *
101: em_est44: *
102: gb_est57: *
103: gb_est58: *
104: em_est45: *
105: gb_est59: *
106: gb_est60: *
107: gb_est61: *
108: gb_est62: *
109: gb_est63: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
C 1	15.4	85.6	77	40	D43091	D43091 D43091 Rice
C 2	14.8	82.2	64	50	F30799	F30799 HSPD21447 H
C 3	14	77.8	94	49	F24375	F24375 HSPD10647 H
C 4	13.8	76.7	58	45	A1343303	A1343303 tbr2909.x
C 5	13.8	76.7	82	38	AA748306	AA748306 oas4a10.s
C 6	13.4	74.4	46	34	AA511536	AA511536 vj24e05.x
C 7	13.4	74.4	74	63	A1971085	A1971085 wr32906.x
C 8	13.4	74.4	77	20	D20607	D20607 HUMG501582
C 9	13.4	74.4	79	29	AA160831	AA160831 z662f02.s
C 10	13.4	74.4	89	44	A1320114	A1320114 b9g05mm.f
C 11	13.2	73.3	76	60	A1789049	A1789049 uk50d02.y
C 12	13.2	73.3	62	62	A1924620	A1924620 wn57d08.x
C 13	13.2	73.3	95	37	AA680947	AA680947 lmfFrm055
C 14	13.2	73.3	97	37	AA680946	AA680946 lmfFrm055
C 15	13.2	73.3	100	50	A1682938	A1682938 tw47d01.x
C 16	13	72.2	25	40	A1000095	A1000095 os61a08.s
C 17	13	72.2	74	37	AA680772	AA680772 lmfFrm030
C 18	13	72.2	99	35	AA591720	AA591720 v148d01.x
C 19	12.8	71.1	46	63	A1957210	A1957210 ul77a04.x
C 20	12.8	71.1	87	94	A0072951	A0072951 EP(2)2165
C 21	12.4	68.9	52	21	T92783	T92783 ye27b02.s1
C 22	12.4	68.9	70	62	A1892436	A1892436 mm39e04.y
C 23	12.4	68.9	76	21	R15882	R15882 ya47a12.r1
C 24	12.4	68.9	87	43	A1161047	A1161047 qc93d02.x
C 25	12.4	68.9	91	40	AA948011	AA948011 oq58906.s
C 26	12.2	67.8	38	48	A1590364	A1590364 ts09e12.x
C 27	12.2	67.8	34	44	A1287522	A1287522 qv86g04.x
C 28	12.2	67.8	37	41	A1018534	A1018534 ou24h12.x
C 29	12.2	67.8	37	47	A1475986	A1475986 l196c09.x
C 30	12.2	67.8	46	23	H24541	H24541 y130f06.r1
C 31	12.2	67.8	46	41	A1033350	A1033350 ox02e11.s
C 32	12.2	67.8	51	79	AA232755	AA232755 f130a01.x
C 33	12.2	67.8	52	23	H24483	H24483 y130a09.r1
C 34	12.2	67.8	60	24	D69796	D69796 CELK092C6F
C 35	12.2	67.8	73	49	A1635197	A1635197 tz22d03.x
C 36	12.2	67.8	74	40	AA984916	AA984916 am63a09.s
C 37	12.2	67.8	76	26	W98562	W98562 mg14f02.r1
C 38	12.2	67.8	76	43	A1194156	A1194156 ue78f02.r
C 39	12.2	67.8	80	44	A1267776	A1267776 aq34a07.x
C 40	12.2	67.8	80	94	AA073477	AA073477 EP(2)2306
C 41	12.2	67.8	80	94	AA073491	AA073491 EP(2)2322
C 42	12.2	67.8	80	94	AA073569	AA073569 EP(2)2416
C 43	12.2	67.8	87	92	AA0911807	AA0911807 LMAJFV1.1
C 44	12.2	67.8	88	45	A1354446	A1354446 qu17f10.x
C 45	12.2	67.8	90	22	R40179	R40179 yf70e09.s1

ALIGNMENTS

RESULT 1

LOCUS D43091 77 bp mRNA EST 04-MAY-1998
 DEFINITION D43091 Rice callus cDNA (H.Uchimiya) Oryza sativa cDNA clone AD544,
 mRNA sequence.

ACCESSION D43091
 VERSION D43091.1 GI:3107351

KEYWORDS EST.
 SOURCE Oryza sativa.

ORGANISM

Oryza sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 eumhyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;

Poaceae; Oryza.

1 (bases 1 to 77)

Uchimiya, H.

On nucleotide sequence of Oryza sativa

Unpublished (1994)

On Jan 14, 1998 this sequence version replaced gi:2754210.

Contact: Hirofumi Uchimiya

Institute of Mol. & Cell. Bioscience, Department of Cellular
 Function

The University of Tokyo
 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113, Japan
 Tel: 03-3812-2111 (ex.7844)
 Fax: 03-3812-2910
 Email: uchimiya@tansei.cc.u-tokyo.ac.jp
 PROJECT= uchimiya.

FEATURES

source
 1..77
 /organism="Oryza sativa"
 /db_xref="taxon:4530"
 /clone="AD544"
 /clone_lib="Rice callus cDNA (H.Uchimiya)"
 /tissue_type="callus"
 BASE COUNT 22 a 25 c 9 g 14 t 7 others

ORIGIN

Query Match 85.6%; Score 15.4; DB 40; Length 77;
 Best Local Similarity 94.1%; Pred. No. 2.2e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ggtgtgtgtgtgtgtt 17
 |||||
 Db 61 GGTGTGTGTGTGT 45

RESULT 2

LOCUS F30799 64 bp mRNA EST 13-MAY-1999
 DEFINITION HSPD21447 HM3 Homo sapiens cDNA clone s400088F12, mRNA sequence.

ACCESSION F30799
 VERSION F30799.1 GI:4816425

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

1 (bases 1 to 64)
 Landfranchi, G., Muraro, F., Caldara, F., Pacchioni, B., Pallavicini, A.,
 Pandolfo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.

Identification of 4370 expressed sequence tags from a
 3'-end-specific cDNA library of human skeletal muscle by DNA
 sequencing and filter hybridization

Genome Res. 6 (1), 35-42 (1996)

JOURNAL

MEDLINE

96276048
 On May 18, 1998 this sequence version replaced gi:3137194.

COMMENT

Contact: Valle G.

CIRIBI Biotechnology Centre

University of Padua

Via Trieste 75, 35121 Padua, Italy

ABI Chromatograms and other information are available on WWW at
<http://group.bio.unipad.it>.

FEATURES

source

Location/Qualifiers

1..64

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="s400088F12"

/clone_lib="HM3"

/sex="female"

/tissue_type="pectoral muscle (after mastectomy)"

/note="Vector: pCDNA1 (Invitrogen); Site_1: BstXI;

Site_2: NotI; The library was constructed by G.

Landfranchi. This library is not subtracted nor normalized.

The first strand cDNA was primed with a biotinylated

oligo-dT-NotI primer

(5'-biotin-AACCGGCTGAGCGCCGCTTTTCTTTTCTTTT-3'). The

ds cDNA was sonicated and size-selected in the range

350-550 bp. The 3' specific fragments were selected by

streptavidin coated magnetic beads, ligated to

non-palindromic BstXI adapters, NotI digested and

directionally cloned into BstXI-NotI cut pCDNA1 vector."

BASE COUNT

7 a 15 c 24 g 18 t

ORIGIN

Query Match 82.2%; Score 14.8; DB 50; Length 64;
Best Local Similarity 88.9%; Pred. No. 3.9e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgctt 18
||||| |||||||
Db 10 GGTGTGAACTTGCTTT 27

RESULT 3
F24375 94 bp mRNA EST 13-MAY-1999
LOCUS HSPD10647 HM3 Homo sapiens cDNA clone s4000008G11, mRNA sequence.
DEFINITION F24375
ACCESSION F24375
VERSION F24375.1 GI:4810001
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 94)
Lanfranchi, G., Muraro, T., Caldara, F., Pacchioni, B., Pallavicini, A.,
Pandolfo, D., Topo, S., Trevisan, S., Scarso, S. and Valle, G.
Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization
Genome Res. 6 (1), 35-42 (1996)
96276048
JOURNAL On Feb 18, 1999 this sequence version replaced gi:4314769.
MEDLINE CRIPI Biotechnology Centre
COMMENT University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
http://grup.bio.unipd.it

FEATURES
source
1..94
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s4000008G11"
/clone_lib="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/note="Vector: pCDNAII (Invitrogen); Site 1: BstXI;
Site 2: NotI; The library was constructed by G.
Lanfranchi. This library is not subcloned nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGCTCGAGCGCGCTTTT-TTTT-TTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT 27 a 28 c 29 g 10 t

Query Match 77.8%; Score 14; DB 49; Length 94;
Best Local Similarity 100.0%; Pred. No. 9e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 gttgtgtgtgctt 17
||||| |||||||
Db 14 GTTGTGCTTGCTT 1

RESULT 4
AI343303 58 bp mRNA EST 08-APR-1999
LOCUS AI343303/c
DEFINITION tb92909.x1 NCI_CGAP_Lu25 Homo sapiens cDNA clone IMAGE:2061856 3'

similar to TR:Q33578 Q33578 KINETOPLAST CR5 ;, mRNA sequence.

ACCESSION AI343303
VERSION AI343303.1 GI:4080509
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 58)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL On Jan 17, 1998 this sequence version replaced gi:2150355.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: I.M.A.G.E. Consortium, LNL.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bdrp/image/image.html

FEATURES
source
1..58
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2061856"
/clone_lib="NCI_CGAP_Lu25"
/tissue_type="bronchialalveolar carcinoma"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: lung; Vector: PAMP1; mRNA made from lung
carcinoma tissue, cDNA made by oligo-dT priming.
Directionally cloned. Size-selected on agarose gel,
average insert size 500 bp. Primary library,
non-amplified."

BASE COUNT 34 a 15 c 8 g 1 t

Query Match 76.7%; Score 13.8; DB 45; Length 58;
Best Local Similarity 88.2%; Pred. No. 1e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 gttgtgtgtgctt 18
||||| |||||||
Db 31 GTTGTGCTTGCTT 15

RESULT 5
AA748306 82 bp mRNA EST 23-JAN-1998
LOCUS aa54a10.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1308762 3'
DEFINITION similar to WP:G36B1.12 CE05379 ;, mRNA sequence.
ACCESSION AA748306
VERSION AA748306.1 GI:2788264
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 82)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL On Jan 14, 1998 this sequence version replaced gi:1798074.
COMMENT Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbp/image/image.html

Trace considered overall poor quality
 Insert Length: 438 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

1. 82
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1308762"
 /clone_lib="NCI CGAP, GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 /note="Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTACCAATCTGAAAGTGAGCGCGCCGCAATTTTCTTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 49 a 22 c 10 g 1 t
 ORIGIN
 Query Match 76.7%; Score 13.8; DB 38; Length 82;
 Best Local Similarity 88.2%; Pred. No. 1.1e+04;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgtt 17
 ||||| ||||| |||||
 DB 23 GGTGTTGTTGCTT 7

RESULT 6
 AA511536 46 bp mRNA EST 08-JUL-1997
 LOCUS vj24e05.r1 Soares mouse NBMH Mus musculus cDNA clone IMAGE:922688
 DEFINITION 5' similar to SW:NM_000502 P03893 NADH-UBIQUINONE OXIDOREDUCTASE
 CHAIN 2 ; mRNA sequence.

ACCESSION AA511536
 VERSION AA511536.1 GI:2249390
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 46)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 The WashU-HHMT Mouse EST Project
 Unpublished (1996)

TITLE The WashU-HHMT Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1395342.
 Contact: Marra M/Mouse EST Project

WashU-HHMT Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (Info@image.llnl.gov) for further information.
 MGI:534904

Trace considered overall poor quality
 Possible reversed clone; similarity on wrong strand
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

1. 46
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:922688"
 /clone_lib="Soares mouse NBMH"
 /sex="male"
 /tissue_type="heart"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTACCAATCTGAAAGTGAGCGCGCCGCAATTTTCTTTT-3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T73 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 15 a 16 c 5 g 10 t
 ORIGIN
 Query Match 74.4%; Score 13.4; DB 34; Length 46;
 Best Local Similarity 93.3%; Pred. No. 1.5e+04;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 gttgtgtgtgtgtt 18
 ||||| ||||| |||||
 DB 43 GTTGTTGTTGCTAT 29

RESULT 7
 A1971085 74 bp mRNA EST 25-AUG-1999
 LOCUS w23906.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:2488570 3'
 DEFINITION similar to gb:M62402 INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 6
 PRECURSOR (HUMAN); mRNA sequence.

ACCESSION A1971085
 VERSION A1971085.1 GI:5767911
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 74)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On Dec 20, 1995 this sequence version replaced gi:1130903.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

TITLE The WashU-HHMT Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1395342.
 Contact: Marra M/Mouse EST Project

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbfp/image/image.html

Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES
source
1.74

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:248570"
/clone_lib="NCI_CGAP_Pr28"
/sex="male"
/dev_stage="adult"
/lab_host="PH10B"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI_CGAP_Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneids
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
20 a 23 c 19 g 12 t

ORIGIN

Query Match 74.4%; Score 13.4; DB 63; Length 74;
Best Local Similarity 93.3%; Pred. No. 1.6e+04;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 gtcgtgtgtgtgct 16
|||||
Db 38 GTCGTGTGTGTGCT 24

RESULT 8
LOCUS D20607 77 bp mRNA EST 30-JUL-1996
DEFINITION HUMS01582 Human promyelocyte Homo sapiens cDNA clone pm152.3',
mRNA sequence.
ACCESSION D20607
VERSION D20607.1 GI:501703
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 77)
Okubo,K., Fukushima,A., Yoshii,J., Niiyama,T., Kojima,Y.,
Yoshinari,H., Arimoto,J. and Matsubara,K.

Gene expression of human promyelocytic cell line HL60 before and
after induction of differentiation. A new application of 3'directed
cDNA sequencing
Unpublished (1993)

CONTACT: Okubo,K., Fukushima,A., Yoshii,J., Niiyama,T., Kojima,Y.,
Yoshinari,H., Arimoto,J. and Matsubara,K.
Institute for Molecular and Cellular Biology
Osaka University
3-1 Yamada-Oka,Suita,Osaka 565,Japan.

FEATURES
source
1.77

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pm152"
/clone_lib="Human promyelocyte"
/note="Female, adult, cell_line = HL60, cell_type =
promyelocyte."

BASE COUNT
27 a 22 c 10 g 17 t 1 others

ORIGIN

Query Match 74.4%; Score 13.4; DB 20; Length 77;
Best Local Similarity 87.5%; Pred. No. 1.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtcgtgtgtgtgct 16
|||||
Db 41 GGNCTTGCTGTGGCT 26

RESULT 9
LOCUS AA160831 79 bp mRNA EST 16-DEC-1996
DEFINITION Z062f02.s1 Stratiagene pancreas (#937208) Homo sapiens cDNA clone
IMAGE:591483 3' similar to SW:RS5_HUMAN P46782 40S RIBOSOMAL
PROTEIN S5.; mRNA sequence.

ACCESSION AA160831
VERSION AA160831.1 GI:1736267
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 79)
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiappelli,B.,
Chisoe,S., Dietrich,N., Dubuque,T., Favellio,A., Gish,N.,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rikfin,L.,
Rohlfing,T., Scheilenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
Trevasakis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
and Marra,M.

Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
On May 18, 1995 this sequence version replaced gi:811261.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 1.

FEATURES
source
1.79

/organism="Homo sapiens"
/db_xref="GDB:462286"
/db_xref="taxon:9606"
/clone="IMAGE:591483"
/clone_lib="Stratiagene pancreas (#937208)"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: pancreas; Vector: pBluescript SK-; Site 1:
ECORI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Pancreatic adenocarcinoma cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
sequence: 5' GAATTCGGCAGCGAG 3' -3' adaptor sequence: 5'
CTCGAGTCTTTTCTTTTCTTTT 3'."

BASE COUNT
16 a 25 c 21 g 13 t 4 others

ORIGIN

Query Match 74.4%; Score 13.4; DB 29; Length 79;
Best Local Similarity 87.5%; Pred. No. 1.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 tgcgtgtgtgtgctt 18
|||||
Db 49 TGTTCGTGATCGCMTT 64

ORIGIN

RESULT	10
LOCUS	A1320114/c
DEFINITION	b9g05mm.f1 Neurospora crassa morning cDNA library Neurospora crassa
ACCESSION	A1320114
VERSION	A1320114
KEYWORDS	cDNA clone b9g05nm 5', mRNA sequence.
SOURCE	A1320114.1 GI:4036096
ORGANISM	EST. Neurospora crassa. Neurospora crassa Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.
REFERENCE	1 (bases 1 to 89)
AUTHORS	Zhu, H., Lai, H., Kupfer, D., Dunlap, J.C. and Roe, B.A.
TITLE	The Neurospora crassa EST databases
JOURNAL	Unpublished (1998)
COMMENT	Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu Department of Chemistry and Biochemistry Advanced Center for Genome Technology, University of Oklahoma 620 Parrington Oval, Norman, OK 73019, USA Tel: 405 325 4912 Fax: 405 325 7762 Email: broe@ou.edu We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center Seq primer: Universal Forward Primer High quality sequence stop: 78. Location/Qualifiers 1..89 /organism="Neurospora crassa" /strain="bd, frq7A" /db_xref="taxon:5141" /clone="b9g05nm" /clone_id="Neurospora crassa morning cDNA library" /tissue_type="tissue harvested following 22hr growth in dark" /note="Vector: pBluescript SK-. Site_1: XbaI; Site_2: EcoRI; See: Bell-Pedersen, D., et al. PNAS 93:1096,1996. 5' end of cDNA cloned into XbaI site of pBluescript; 3' end of cDNA cloned into EcoRI site of pBluescript"
BASE COUNT	20 a 35 c 15 g 19 t
ORIGIN	
Query Match	74.4%; Score 13.4; DB 44; Length 89;
Best Local Similarity	93.3%; Pred. No. 1.6e+04;
Matches	14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	1 gggttggtgtgtgc 15 Db 21 GGGTAGCTGTTGC 7
RESULT	11
LOCUS	A1789049/c
DEFINITION	u550d02.y1 Sugano mouse kidney mKia Mus musculus cDNA clone
IMAGE:	1972419 5' similar to TR:Q28170 Q28170 POLYUBQUITIN ; , mRNA sequence.
ACCESSION	A1789049
VERSION	A1789049.1
KEYWORDS	GI:5336765
SOURCE	EST. house mouse. Mus musculus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 76) Marras, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,

TITLE	Waterston,R. and Wilson,R					
JOURNAL	The Washn-NCI Mouse EST Project 1999					
COMMENT	Unpublished (1999) On Dec 20, 1995 this sequence version replaced gi:1135238. Contact: Marra M/Washu-NCI Mouse EST Project 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouse@waterston.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (Info@image.llnl.gov) for further information. MGI:989159					
FEATURES	Trace considered overall poor quality Possible reversed clone; similarity on wrong strand Seq primer: custom primer used High quality sequence stop: 1.					
SOURCE	Location/Qualifiers 1..76 /organism="Mus musculus" /strain="C57BL" /db_xref="taxon:10090" /clone="IMAGE:1972419" /clone_1lb="Sugano mouse kidney mkia" /sex="female" /dev_stage="adult" /lab_host="DH10B" /note="Organ: kidney; Vector: PME18S-FL3; Site_1: DraIII (CACGCTGTG); Site_2: DraIII (CACCAAGTGG); 1st strand cDNA was primed with an oligo(dT) primer [ATGCGCCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TCGTGGCCACTACGG], digested and cloned into distinct DraIII sites of the PME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCAGTGG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTTGCTCTTAAGAAGCTGCC and 3' end primer CGACCTGACGCTCAGCACA."					
BASE COUNT	17 a 19 c 19 g 21 t					
ORIGIN						
Query Match	73.3%; Score 13.2; DB 60; Length 76;					
Best Local Similarity	83.3%; Pred. No. 1.ge+04;					
Matches 15; Conservative	0; Mismatches 3; Indels 0; Gaps 0					
Oy	1 ggttgtgtttgctt 18 Db 73 GGATGATGGTGTGCCTGT 56					
RESULT 12						
AI924620/c						
LOCUS	AI924620 76 bp mRNA EST 02-SEP-1999					
DEFINITION	wn57d08.x1 NCI-CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2449551 3'					
ACCESSION	AI924620					
VERSION	similar to gb:M60858_rnal NUCLEOLIN (HUMAN);, mRNA sequence.					
KEYWORDS	AI924620.1 GI:5660584					
SOURCE	EST.					
ORGANISM	human.					
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.					
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)					
JOURNAL	On Dec 20, 1995 this sequence version replaced gi:1135542.					
COMMENT	Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550					

Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLINL at:
www.bio.lnlnl.gov/dbp/image/image.html

Trace considered overall poor quality
Seq primer: -40UP from Gabc
High quality sequence stop: 1.

FEATURES

source

1. /76
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2449551"
/clone_lib="NCI CGAP_Lu19"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from a
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 28 a 24 c 16 g 8 t
ORIGIN

Query Match 73.3%; Score 13.2; DB 62; Length 76;
Best Local Similarity 83.3%; Pred. No. 1.9e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgtt 18
||||| ||| |||
Db 65 GGTGTGTCTCTGCTTT 48

RESULT 13

AA680947/c

LOCUS AA680947 95 bp mRNA EST 09-DEC-1998
DEFINITION LmfrAm0555 Leishmania major Amastigote Lambda Zap II library
Leishmania major cDNA clone R69 5', mRNA sequence.

ACCESSION AA680947
VERSION AA680947.1 GI:2662952
KEYWORDS EST.
SOURCE Leishmania major.
ORGANISM Leishmania major
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Leishmania.
1 (bases 1 to 95)

REFERENCE 1 (bases 1 to 95)
AUTHORS Norrish,A.R., Dyall,S.D., Smith,D.F. and Blackwell,J.M.
TITLE Analysis of Leishmania Major Amastigote Expressed Sequence Tags
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1997 this sequence version replaced g1:2264540.
Contact: Blackwell JM
Wellcome Trust/MRC Building, Addenbrooke's Hospital, Hills Road,
Cambridge CB2 2XY, UK
Tel: 01223 336 143
Fax: 01223 331 206
Email: jmb37@cus.cam.ac.uk
PCR Primers
FORWARD: GTAAACGACGACGACG
BACKWARD: GGAACAGCTATGACCATG

Seq primer: AATTAACTTCACATAAGG
High quality sequence stop: 95.

FEATURES

source

1. /95
/organism="Leishmania major"
/strain="Friedlin"
/db_xref="taxon:5664"
/clone="R69"
/clone_lib="Leishmania major Amastigote Lambda Zap II
library"
/cell_type="Amastigote"
/note="Vector: Lambda Zap II; Site_1: XhoI; Site_2: NotI"

BASE COUNT 54 a 39 c 2 g 0 t
ORIGIN

Query Match 73.3%; Score 13.2; DB 37; Length 95;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgtt 18
||||| ||| |||
Db 53 GGTGTGTGTGTGTGT 36

RESULT 14

AA680946/c

LOCUS AA680946 97 bp mRNA EST 09-DEC-1998
DEFINITION LmfrAm0554 Leishmania major Amastigote Lambda Zap II library
Leishmania major cDNA clone R68 5', mRNA sequence.

ACCESSION AA680946
VERSION AA680946.1 GI:2662951
KEYWORDS EST.
SOURCE Leishmania major.
ORGANISM Leishmania major
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Leishmania.
1 (bases 1 to 97)

REFERENCE 1 (bases 1 to 97)
AUTHORS Norrish,A.R., Dyall,S.D., Smith,D.F. and Blackwell,J.M.
TITLE Analysis of Leishmania Major Amastigote Expressed Sequence Tags
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1997 this sequence version replaced g1:2264539.
Contact: Blackwell JM
Wellcome Trust/MRC Building, Addenbrooke's Hospital, Hills Road,
Cambridge CB2 2XY, UK
Tel: 01223 336 143
Fax: 01223 331 206
Email: jmb37@cus.cam.ac.uk
PCR Primers
FORWARD: GTAAACGACGACGACG
BACKWARD: GGAACAGCTATGACCATG
Seq primer: AATTAACTTCACATAAGG
High quality sequence stop: 97.

FEATURES
source
1. /97
/organism="Leishmania major"
/strain="Friedlin"
/db_xref="taxon:5664"
/clone="R68"
/clone_lib="Leishmania major Amastigote Lambda Zap II
library"
/cell_type="Amastigote"
/note="Vector: Lambda Zap II; Site_1: XhoI; Site_2: NotI"

BASE COUNT 56 a 39 c 2 g 0 t
ORIGIN

Query Match 73.3%; Score 13.2; DB 37; Length 97;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgtt 18

Db 53 GGCTTGTGTGTGGTGT 36

RESULT 15

AI682938/C

LOCUS

AI682938 100 bp mRNA EST 15-DEC-1999
 tw47d01.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2262817 3'
 similar to TR:Q91236 Q91236 ZONA PELUCIDA PROTEIN ; contains
 MER22.b2 MSRI repetitive element ; , mRNA sequence.

ACCESSION

AI682938
 AI682938.1 GI:4893120

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

REFERENCE

1 (bases 1 to 100)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL

Unpublished (1997)
 On May 18, 1998 this sequence version replaced gi:3137345.

COMMENT

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/BLNT at:

www.bio.lnl.gov/bhrp/image/image.html

Trace considered overall poor quality

Insert Length: 1340 Std Error: 0.00

Seq Primer: -400P from Gibco

High quality sequence stop: 1.

FEATURES

source

1..100

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2262817"

/clone_lib="NCI-CGAP_Ut1"

/tissue_type="well-differentiated endometrial

adenocarcinoma, 7 pooled tumors"

/lab_host="DH10B"

/note="Organ: uterus; Vector: PCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.75 kb. Life Technologies catalog #:

11538-014"

BASE COUNT

33 a 47 c 20 g 0 t

ORIGIN

Query Match

Best Local Similarity 73.3%; Score 13.2; DB 50; Length 100;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgtgttt 18

DB 86 GGCTTGTGTGTGGTGT 69

Search completed: June 4, 2000, 13:53:19
 Job time: 20943 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:35 ; Search time 1236.38 Seconds
(without alignments)
-14.163 Million cell updates/sec

Title: US-09-164-714-13
Perfect score: 18
Sequence: 1 ggtgtgtgtgtgtgttt 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_bal:*
2: gb_bal:*
3: gb_bal:*
4: gb_bal:*
5: gb_bal:*
6: gb_bal:*
7: gb_bal:*
8: gb_bal:*
9: gb_bal:*
10: gb_bal:*
11: gb_bal:*
12: gb_bal:*
13: gb_bal:*
14: gb_bal:*
15: gb_bal:*
16: gb_bal:*
17: gb_bal:*
18: gb_bal:*
19: gb_bal:*
20: gb_bal:*
21: gb_bal:*
22: gb_bal:*
23: gb_bal:*
24: gb_bal:*
25: gb_bal:*
26: gb_bal:*
27: gb_bal:*
28: gb_bal:*
29: gb_bal:*
30: gb_bal:*
31: gb_bal:*
32: gb_bal:*
33: gb_bal:*
34: gb_bal:*
35: gb_bal:*
36: gb_bal:*
37: gb_bal:*
38: gb_bal:*
39: gb_bal:*
40: gb_bal:*
41: gb_bal:*
42: gb_bal:*
43: gb_bal:*
44: gb_bal:*

45: gb_hcg7:*
46: em_hcg1:*
47: em_hcg2:*
48: em_hcg3:*
49: em_hcg4:*
50: gb_pl3:*
51: gb_pl5:*
52: gb_hcg8:*
53: gb_hcg9:*
54: gb_hcg10:*
55: gb_hcg11:*
56: gb_hcg12:*
57: gb_hcg13:*
58: gb_hcg14:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14.4	80.0	90	5	I36132	Sequence 16
2	13.8	76.7	36	5	I27773	Sequence 5
3	13.8	76.7	36	5	I27775	Sequence 7
4	13.2	73.3	20	5	I76185	Sequence 4
5	13.2	73.3	24	5	I27784	Sequence 16
6	13.2	73.3	41	5	A51237	Sequence 8
7	13.2	73.3	45	5	A77178	Sequence 6
8	13.2	73.3	45	5	A77180	Sequence 6
9	13.2	73.3	45	5	I20213	Sequence 8
10	13.2	73.3	45	5	I20215	Sequence 6
11	13.2	73.3	51	5	A12323	Sequence 8
12	13.2	73.3	51	5	A12324	Sequence 8
13	13.2	73.3	51	5	A12326	Sequence 8
14	13.2	73.3	51	5	A12327	Sequence 8
15	13.2	73.3	92	12	D55656S10	Sequence 16
16	12.8	71.1	26	5	AR018921	Sequence 4
17	12.8	71.1	26	5	AR066549	Sequence 43
18	12.8	71.1	26	5	I72266	Sequence 43
19	12.8	71.1	26	5	I76833	Sequence 43
20	12.8	71.1	66	5	A02001	Artificial
21	12.8	71.1	66	5	A06447	Artificial
22	12.8	71.1	66	5	AF027108	USC11ago
23	12.8	71.1	76	5	E02399	DNA encodin
24	12.8	71.1	77	5	E01808	DNA encodin
25	12.8	71.1	77	5	E02146	DNA encodin
26	12.8	71.1	77	5	E02398	DNA encodin
27	12.8	71.1	93	5	A22901	H. insoliens
28	12.8	71.1	93	5	E03016	DNA encodin
29	12.4	68.9	54	5	I50003	Sequence 1
30	12.4	68.9	62	5	I50012	Sequence 10
31	12.4	68.9	63	5	I50013	Sequence 11
32	12.4	68.9	85	5	AR051981	Sequence 65
33	12.4	68.9	85	5	I52155	Sequence 65
34	12.4	68.9	90	9	HSDPBVBRB	Y15719 Homo sapien
35	12.2	67.8	26	5	I26256	Sequence 3
36	12.2	67.8	40	5	AR053640	Sequence 3
37	12.2	67.8	40	5	AR053647	Sequence 3
38	12.2	67.8	40	5	I86252	Sequence 6
39	12.2	67.8	40	5	I86255	Sequence 9
40	12.2	67.8	78	5	A37066	Sequence 9
41	12.2	67.8	78	5	A37067	Sequence 10
42	12.2	67.8	83	13	AU025339	Rattus no
43	12.2	67.8	94	13	G38864	Plasmo
44	12.2	67.8	98	5	I91506	Sequence 40
45	12	66.7	99	10	HSPHA825	Human mRNA

ALIGNMENTS

RESULT 1
LOCUS I36132 90 bp DNA PAT 13-MAY-1997
DEFINITION Sequence 16 from patent US 5604131.
ACCESSION I36132
VERSION I36132.1 GI:2087356
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 90)
AUTHORS Wadsworth,S., Snyder,B., Reddy,V.B. and Wei,C.
TITLE CDNA-genomic DNA hybrid sequence encoding APP770 containing a
JOURNAL genomic DNA insert of the KI and OX-2 regions
FEATURES
Source Location/Qualifiers
1..90
/organism="unknown"
BASE COUNT 18 a 20 c 12 g 40 t
ORIGIN

Query Match 80.0%; Score 14.4; DB 5; Length 90;
Best Local Similarity 93.8%; Pred. No. 2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 tttgtgtgtgtgtgt 18
|||||
DB 33 tttgtgtgtgtgtgt 48

RESULT 2
LOCUS I27773 36 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 5 from patent US 5567604.
ACCESSION I27773
VERSION I27773.1 GI:1818549
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 36)
AUTHORS Rando,R.F., Fennewald,S., Zendegeul,J.G. and OJwang,J.O.
TITLE Anti-viral guanosine-rich oligonucleotides
JOURNAL Patent: US 5567604-A 5 22-OCT-1996;
FEATURES
Source Location/Qualifiers
1..36
/organism="unknown"
BASE COUNT 0 a 0 c 24 g 12 t
ORIGIN

Query Match 76.7%; Score 13.8; DB 5; Length 36;
Best Local Similarity 88.2%; Pred. No. 3.8e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgt 17
|||||
DB 9 ggtgtgtgtgtgtgt 25

RESULT 3
LOCUS I27775 36 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 7 from patent US 5567604.
ACCESSION I27775
VERSION I27775.1 GI:1818551
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 36)

AUTHORS Rando,R.F., Fennewald,S., Zendegeul,J.G. and OJwang,J.O.
TITLE Anti-viral guanosine-rich oligonucleotides
JOURNAL Patent: US 5567604-A 7 22-OCT-1996;
FEATURES
Source Location/Qualifiers
1..36
/organism="unknown"
BASE COUNT 0 a 0 c 24 g 12 t
ORIGIN

Query Match 76.7%; Score 13.8; DB 5; Length 36;
Best Local Similarity 88.2%; Pred. No. 3.8e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgt 17
|||||
DB 9 ggtgtgtgtgtgtgt 25

RESULT 4
LOCUS I76185 20 bp DNA PAT 03-APR-1998
DEFINITION Sequence 4 from patent US 5691145.
ACCESSION I76185
VERSION I76185.1 GI:3012339
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 20)
AUTHORS Plitner,J.Bruce, Vonk,G.P. and Nadeau,J.G.
TITLE Detection of nucleic acids using G-quartets
JOURNAL Patent: US 5691145-A 4 25-NOV-1997;
FEATURES
Source Location/Qualifiers
1..20
/organism="unknown"
BASE COUNT 0 a 0 c 8 g 12 t
ORIGIN

Query Match 73.3%; Score 13.2; DB 5; Length 20;
Best Local Similarity 83.3%; Pred. No. 7.6e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgt 18
|||||
DB 1 ggtgtgtgtgtgtgt 18

RESULT 5
LOCUS I27784 24 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 16 from patent US 5567604.
ACCESSION I27784
VERSION I27784.1 GI:1818560
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 24)
AUTHORS Rando,R.F., Fennewald,S., Zendegeul,J.G. and OJwang,J.O.
TITLE Anti-viral guanosine-rich oligonucleotides
JOURNAL Patent: US 5567604-A 16 22-OCT-1996;
FEATURES
Source Location/Qualifiers
1..24
/organism="unknown"
BASE COUNT 0 a 0 c 14 g 10 t
ORIGIN

Query Match 73.3%; Score 13.2; DB 5; Length 24;
Best Local Similarity 83.3%; Pred. No. 7.7e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

ACCESSION I20215
VERSION I20215.1 GI:1600570
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 45)
AUTHORS Kriebber, K., Moroney, S., Pluckthun, A. and Schneider, C.
TITLE Method for in vivo selection of ligand-binding proteins
JOURNAL Patent: US 5514548-A 8 07-MAY-1996;
FEATURES Location/Qualifiers
source 1..45
/organism="unknown"
BASE COUNT 9 a 7 c 13 g 16 t
ORIGIN

Query Match 73.3%; Score 13.2; DB 5; Length 45;
Best Local Similarity 83.3%; Pred. No. 8e+03; 3; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgctt 18
||||| ||||| ||| |||
Db 15 GGTGATGCTGATGATT 32

RESULT 11
LOCUS A12323 51 bp DNA PAT 06-DEC-1993
DEFINITION oligonucleotide.
ACCESSION A12323
VERSION A12323.1 GI:491330
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 51)
ARTIFICIAL SEQUENCE.
AUTHORS
TITLE HYBRID PROTEINS OR POLYPEPTIDES
JOURNAL Patent: WO 8802757-A 24 21-APR-1988;
FEATURES Location/Qualifiers
source 1..51
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 15 a 27 c 7 g 2 t
ORIGIN

Query Match 73.3%; Score 13.2; DB 5; Length 51;
Best Local Similarity 83.3%; Pred. No. 8.1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgctt 18
||||| ||||| ||||| |||
Db 39 GCGGTGGGGTGGCGTT 22

RESULT 12
LOCUS A12324 51 bp DNA PAT 06-DEC-1993
DEFINITION oligonucleotide.
ACCESSION A12324
VERSION A12324.1 GI:489519
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 51)
ARTIFICIAL SEQUENCE.
AUTHORS
TITLE HYBRID PROTEINS OR POLYPEPTIDES
JOURNAL Patent: WO 8802757-A 25 21-APR-1988;
FEATURES Location/Qualifiers
source 1..51

/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 2 a 7 c 27 g 15 t
ORIGIN

Query Match 73.3%; Score 13.2; DB 5; Length 51;
Best Local Similarity 83.3%; Pred. No. 8.1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgctt 18
||||| ||||| ||||| |||
Db 13 GCGGTGGGGTGGCGTT 30

RESULT 13
LOCUS A12596 51 bp DNA PAT 05-JAN-1994
DEFINITION oligonucleotide.
ACCESSION A12596
VERSION A12596.1 GI:491421
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 51)
ARTIFICIAL SEQUENCE.
AUTHORS
TITLE RECOMBINANT VIRUS
JOURNAL Patent: WO 8701386-A 12 12-MAR-1987;
FEATURES Location/Qualifiers
source 1..51
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 15 a 27 c 7 g 2 t
ORIGIN

Query Match 73.3%; Score 13.2; DB 5; Length 51;
Best Local Similarity 83.3%; Pred. No. 8.1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgctt 18
||||| ||||| ||||| |||
Db 39 GCGGTGGGGTGGCGTT 22

RESULT 14
LOCUS A12597 51 bp DNA PAT 05-JAN-1994
DEFINITION oligonucleotide.
ACCESSION A12597
VERSION A12597.1 GI:489543
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 51)
ARTIFICIAL SEQUENCE.
AUTHORS
TITLE RECOMBINANT VIRUS
JOURNAL Patent: WO 8701386-A 13 12-MAR-1987;
FEATURES Location/Qualifiers
source 1..51
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 2 a 7 c 27 g 15 t
ORIGIN

Query Match 73.3%; Score 13.2; DB 5; Length 51;
Best Local Similarity 83.3%; Pred. No. 8.1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgctt 18

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:08 ; Search time 322.35 Seconds
(without alignments)
13.971 Million cell updates/sec

Title: US-09-164-714-12

Perfect score: 18
Sequence: 1 ggcagcaaacaccagctag 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N.Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	13.8	76.7	25	1	063272	Upstream activatio
C 2	13.8	76.7	25	1	081991	Upstream activatio
C 3	13.8	76.7	31	1	063273	Upstream activatio
C 4	13.8	76.7	31	1	063274	Upstream activatio
C 5	13.4	74.4	30	1	054365	Primer/probe D1 sp
C 6	12.8	71.1	22	1	X24317	Rat bone mineralis
C 7	12.4	68.9	30	1	054366	Primer/probe D2 sp
C 8	12.4	68.9	56	1	T25702	Human gene signal
C 9	12.2	67.8	30	1	V16996	Telomerase competi
C 10	12.2	67.8	48	1	012661	SUC2 signal peptid
C 11	12.2	67.8	55	1	034057	Downstream sequenc
C 12	12.2	67.8	57	1	004915	Wild type signal s
C 13	12.2	67.8	57	1	004916	Modified signal se
C 14	12.2	67.8	57	1	004917	Modified signal se
C 15	12.2	67.8	57	1	004918	Modified signal se
C 16	12.2	67.8	57	1	N81440	Modified signal se
C 17	12.2	67.8	63	1	004920	Sequence of yeast
C 18	12.2	67.8	64	1	043329	Modified signal se
C 19	12.2	67.8	66	1	004921	Sequence of PCR pr
C 20	12.2	67.8	69	1	042522	Modified signal se
C 21	12.2	67.8	75	1	N71270	Sequence encoding
C 22	12.2	67.8	75	1	011885	Sequence encoding
C 23	12.2	67.8	75	1	011887	Invertase signal -
C 24	12.2	67.8	75	1	N40005	Invertase signal -
C 25	12.2	67.8	75	1	N40006	Yeast preinvertase
C 26	12.2	67.8	75	1	N40008	Invertase signal s
C 27	12.2	67.8	75	1	083672	Invertase signal s
C 28	12.2	67.8	84	1	083644	Partial IGF-I fusi
C 29	12.2	67.8	84	1	083673	Oligo for fusing t
C 30	12.2	67.8	92	1	056743	Partial IGF-I fusi
C 31	12.2	67.8	94	1	T61041	Sequence of PCR pr
C 32	11.8	65.6	23	1	T17851	Antisense PCR prim
C 33	11.8	65.6	29	1	V91720	Canola napin promc
C 34	11.8	65.6	29	1	V91593	Human C-raf hamme

C 35	11.8	65.6	30	1	054367	Primer/probe D3 sp
C 36	11.8	65.6	30	1	054368	Primer/probe D4 sp
C 37	11.8	65.6	30	1	054369	Primer/probe D5 sp
C 38	11.8	65.6	40	1	030844	Type III procollag
C 39	11.8	65.6	81	1	073051	Conserved sequence
C 40	11.6	64.4	21	1	X15036	Antisense PCR prim
C 41	11.6	64.4	21	1	X15012	Probe used to isol
C 42	11.6	64.4	37	1	049410	PCR primer for amp
C 43	11.6	64.4	42	1	X16930	Primer 028 for con
C 44	11.4	63.3	17	1	V97277	Human EGF-R target
C 45	11.4	63.3	17	1	V97278	Human EGF-R target

ALIGNMENTS

RESULT 1	063272/c	063272 standard; cDNA; 25 BP.
ID	063272	
AC	063272	
DT	14-DEC-1994	(first entry)
DE	Upstream activation sequence of AOX2 promoter.	
KW	AOX2; Alcohol oxygenase; Promoter region; enhanced activity; UAS;	
KW	upstream activation sequence; heterologous protein expression; ss.	
OS	Pichia pastoris.	
PN	EP-595334-A.	
PD	04-MAY-1994.	
PF	28-OCT-1993; 117527.	
PR	30-OCT-1992; JP-293315.	
PR	06-AUG-1993; JP-215306.	
PA	(GREG) GREEN CROSS CORP.	
PI	Chuanji M, Hiramatsu R, Miura M, Ohi H, Ohmura T;	
PI	Uno S, Hiramatsu F;	
DR	WPI: 94-145941/18.	
PT	New mutant AOX promoter - having a deletion, replacement or	
PT	insertion of nucleotide to increase transcription activity in	
PT	prod. of proteins	
PS	Example 11; Page 13; 48pp; English.	
CC	063272 is a putative UAS of AOX2 promoter.	
CC	The promoter is mutated such that it has an enhanced activity. The	
CC	mutant AOX2 promoter is useful in an expression vector allowing	
CC	heterologous protein expression.	
SQ	Sequence 25 BP; 7 A; 3 C; 5 G; 10 T;	
Query Match	76.7%; Score 13.8; DB 1; Length 25;	
Best local similarity	88.2%; Pred. No. 1.1e+02;	
Matches 15; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
Qy	1 ggcagcaaacaccagcta 17	
Db	19 GCGCAAAATAACCTTA 3	
RESULT 2	081991/c	
ID	081991	standard; DNA; 25 BP.
AC	081991	
DT	28-SEP-1995	(first entry)
DE	Upstream activation sequence for mutant alcohol oxidase promoter.	
KW	Alcohol oxidase promoter; upstream activation sequence; ss.	
OS	Synthetic.	
PN	EP-639643-A.	
PD	22-FEB-1995.	
PF	26-JUL-1994; 111628.	
PR	27-JUL-1993; JP-185003.	
PA	(GREG) GREEN CROSS CORP.	
PI	Hiramatsu R, Miura M, Ohi H, Ohmura T;	
DR	WPI: 95-083455/12.	
PT	New mutant alcohol oxidase promoter - for heterologous protein	
PT	expressions esp. in yeast, has at least one added sequence	
PT	enhancing transcriptional activity	
PT	Claim 1; Page 18; 40pp; English.	

CC Q81991 is an upstream activation sequence which is added in the
CC reverse direction or in both the forward and reverse directions at
CC the 5' end of a partial DNA fragment of a wild-type alcohol oxidase
CC (AOX2) promoter. The sequence of the wild-type AOX2 promoter is
CC given in Q81992. The partial fragment of AOX2 is indicated in FT
CC Q81992. The partial fragment also pref. has certain mutations, also
CC indicated in FT. Incorporation of the upstream activation sequence
CC provides more efficient expression of heterologous downstream genes.
SQ Sequence 25 BP; 7 A; 3 C; 5 G; 10 T;

Query Match 76.7%; Score 13.8; DB 1; Length 25;
Best Local Similarity 88.2%; Pred. No. 1.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaacagccta 17
|||||
DB 19 GCGACAAATAGCCTA 3

RESULT 3
063273/c
ID 063273; standard; cDNA; 31 BP.
AC 063273;
DT 15-DEC-1994 (first entry)
DE Upstream activation sequence of AOX2 promoter.
KW AOX2; Alcohol oxygenase; Promoter region; enhanced activity; UAS;
KM upstream activation sequence; heterologous protein expression; ss.
OS Pichia pastoris.
PI Key Location/Qualifiers
FT enhancer 6..31
FT /*tag= a
FT /note= "putative upstream activation sequence"

EP-595334-A.
PD 04-MAY-1994.
PR 28-OCT-1993; 117527.
PR 30-OCT-1992; JP-293315.
PR 06-AUG-1993; JP-215306.
PA (GRNC) GREEN CROSS CORP.
PI Chuganji M, Hiramatsu R, Miura M, Ohl H, Ohmura T;
PI Uno S, Hiramatsu F;
DR WPI: 94-145941/18.
PT New mutant AOX promoter - having a deletion, replacement or
PT insertion of nucleotide to increase transcription activity in
PT prodn. of proteins
PS Example 11; Page 13; 48pp; English.
CC 063272 is a putative UAS of AOX2 promoter from Pichia pastoris. The
CC 25bp sequence was chemically synthesised and EcoRI sites added.
CC The promoter is mutated such that it has an enhanced activity. The
CC mutant AOX2 promoter is useful in an expression vector allowing
CC heterologous protein expression.
SQ Sequence 31 BP; 9 A; 4 C; 6 G; 12 T;

Query Match 76.7%; Score 13.8; DB 1; Length 31;
Best Local Similarity 88.2%; Pred. No. 1.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaacagccta 17
|||||
DB 24 GCGACAAATAGCCTA 8

RESULT 4
063274
ID 063274; standard; cDNA; 31 BP.
AC 063274;
DT 15-DEC-1994 (first entry)
DE Upstream activation sequence of AOX2 promoter.
KW AOX2; Alcohol oxygenase; Promoter region; enhanced activity; UAS;
KM upstream activation sequence; heterologous protein expression; ss.
OS Pichia pastoris.
PI EP-595334-A.

PD 04-MAY-1994.
PR 28-OCT-1993; 117527.
PR 30-OCT-1992; JP-293315.
PR 06-AUG-1993; JP-215306.
PA (GRNC) GREEN CROSS CORP.
PI Chuganji M, Hiramatsu R, Miura M, Ohl H, Ohmura T;
PI Uno S, Hiramatsu F;
DR WPI: 94-145941/18.
PT New mutant AOX promoter - having a deletion, replacement or
PT insertion of nucleotide to increase transcription activity in
PT prodn. of proteins
PS Example 11; Page 13; 48pp; English.
CC 063273 is a putative UAS of AOX2 promoter from Pichia pastoris. The
CC 25bp sequence (reverse strand) was chemically synthesised and EcoRI
CC sites added. The promoter is mutated such that it has an enhanced
CC activity. The mutant AOX2 promoter is useful in an expression vector
CC allowing heterologous protein expression.
SQ Sequence 31 BP; 12 A; 6 C; 4 G; 9 T;

Query Match 76.7%; Score 13.8; DB 1; Length 31;
Best Local Similarity 88.2%; Pred. No. 1.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaacagccta 17
|||||
DB 12 GCGACAAATAGCCTA 28

RESULT 5
054365/c
ID 054365; standard; DNA; 30 BP.
AC 054365;
DT 24-JUN-1994 (first entry)
DE Primer/probe D1 specific for Erwinia chrysanthemi.
KW Erwinia chrysanthemi strain EC16; pectinolytic bacteria; pelE gene;
KM polymerase chain reaction amplification; detection; pectate lyase;
KW plant pathogen; ss.
OS Erwinia chrysanthemi (strain EC16).
PD W09325708-A.
PD 23-DEC-1993.
PR 04-JUN-1993; F00540.
PR 05-JUN-1992; FR-006888.
PA (INRA-) INRA INST NAT AGRONOMIQUE PARIS-GRIGNON.
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
PI Darresse A, Koloujansky A, Bertheau Y;
DR WPI: 94-007565/01.
PT Identification and detection of specific Erwinia species.
PT sub-species and pathogens - by amplifying pectate lyase DNA then
PT hybridisation testing
PS Claim 5; Page 33; 47pp; French.
CC The sequence complementary to nucleotides 672-701 of the pectate
CC lyase gene pelE from E. chrysanthemi strain 16. The oligonucleotide
CC is used with probe/primer C (054362-054364) to specifically amplify
CC a region of the pel gene. Early diagnosis of plant diseases caused
CC by E. chrysanthemi is possible using the primers.
CC See 054371 for the full-length pelE gene.
SQ Sequence 30 BP; 3 A; 6 C; 11 G; 10 T;

Query Match 74.4%; Score 13.4; DB 1; Length 30;
Best Local Similarity 93.3%; Pred. No. 1.8e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 cgacaaacagcct 16
|||||
DB 27 CGACAAACATCCT 13

RESULT 6
X24317/c
ID X24317; standard; DNA; 22 BP.

AC X24317;
DT 07-JUN-1999 (first entry)
DE Rat bone mineralisation protein RIMP reverse PCR primer.
KW RIMP; LMP; LIM mineralisation protein; bone mineralisation; rat;
KW cytokine; spine fusion; fracture repair; bone grafting;
KW osteoporosis; gene therapy; PCR; primer; ss.
OS Synthetic.
OS Rattus norvegicus.
PN W09906563-A1.
PD 11-FEB-1999.
PF 29-JUL-1998; U15814.
PR 02-APR-1998; US-080407.
PR 30-JUL-1997; US-054219.
PA (UYEM-) UNTV EMORY.
PI Boden S, Hair G;
DR WPI; 99-153793/13.
PT LIM mineralisation proteins - used to induce bone formation from
PT osteogenic precursors, e.g. for treating osteoporosis or repairing
PT fractures
PS Example 16: Page 53; 67pp; English.
CC This reverse primer, and a forward primer (see X24316), are based
CC on rat LIM mineralisation protein (RIMP) cDNA (see X24307).
CC The primers were used to amplify a unique 223 bp sequence from
CC rat RIMP cDNA. A similar PCR product (see X24319) was isolated
CC from human MG63 osteosarcoma cell cDNA with the same PCR primers.
CC The human amplified sequence showed 95% homology to the amplified
CC rat sequence. The reverse primer was also used with a different
CC specific primer (see X24318) in a PCR amplification of MG63 cDNA
CC to obtain another partial sequence (see X24320). A full-length
CC human LMP cDNA (see X24322) was subsequently obtained. The
CC encoded protein (see M97844) enhances bone mineralisation in vitro
CC and in vivo. Clinical applications include enhancement of bone
CC repair in fractures, bone defects, bone grafting, and normal
CC homeostasis in patients presenting with osteoporosis. Methods of
CC inducing bone formation using transfected osteogenic precursor
CC cells are provided.
SQ Sequence 22 BP; 1 A; 4 C; 8 G; 9 T;

Query Match 71.1%; Score 12.8; DB 1; Length 22;
Best Local Similarity 87.5%; Pred. No. 3.5e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 gacaaacagcctag 18
DB 19 GACAAAACGACACAG 4

RESULT 7
O54366/c
ID 054366 standard; DNA; 30 BP.
AC 054366;
DT 24-JUN-1994 (first entry)
DE Primer/probe D2 specific for *Erwinia chrysanthemi*.
KW *Erwinia chrysanthemi* strain EC16; pectinolytic bacterium; pelf gene;
KW polymerase chain reaction amplification; detection; pectate lyase;
KW plant pathogen; ss.
OS *Erwinia chrysanthemi* (strain EC16).
PN W09325708-A.
PD 23-DEC-1993.
PF 04-JUN-1993; F00540.
PR 05-JUN-1992; FR-006888.
PA (INRA-) INRA INST NAT AGRONOMIQUE PARIS-GRIGNON.
PA (INRA-) INRA INST NAT RECH AGRONOMIQUE.
PI Darasse A, Kotoujansky A, Bertheau Y;
DR WPI: 94-007565/01.
PT Identification and detection of specific *Erwinia* species,
PT sub-species and pathogens - by amplifying pectate lyase DNA then
PT hybridisation testing
PS Claim 5; Page 33; 47pp; French.
CC Probe/primer D2 is one of 5 preferred oligonucleotides derived from
CC the sequence complementary to nucleotides 672-701 of the pectate
CC lyase gene pelf from *E. chrysanthemi* strain 16. The oligonucleotide

CC is used with probe/primer C (O54362-O54364) to specifically amplify
CC a region of the pelf gene. Early diagnosis of plant diseases caused
CC by *E. chrysanthemi* is possible using the primers.
CC See 054371 for the full-length pelf gene.
SQ Sequence 30 BP; 5 A; 6 C; 9 G; 10 T;

Query Match 68.9%; Score 12.4; DB 1; Length 30;
Best Local Similarity 92.9%; Pred. No. 5.8e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 gacaaacagcct 16
DB 26 GACAAAACGACCTCT 13

RESULT 8
T25702/c
ID T25702 standard; cDNA to mRNA; 56 BP.
AC T25702;
DT 10-OCT-1996 (first entry)
DE Human gene signature HUMGS07903.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PF 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1910; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 56 BP; 13 A; 15 C; 16 G; 12 T;

Query Match 68.9%; Score 12.4; DB 1; Length 56;
Best Local Similarity 92.9%; Pred. No. 6.1e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 caaaacagcctag 18
DB 32 CAAAACGACGCCAG 19

RESULT 9
V16996/c
ID V16996 standard; DNA; 30 BP.
AC V16996;
DT 13-AUG-1998 (first entry)
DE Telomerase competitor DNA oligonucleotide.
KW Human; telomerase reverse transcriptase; hTERT; TERT; diagnosis;

KM prognosis; cell proliferation; cancer; ageing; ribonucleoprotein; ss.
OS Synthetic.
PN GB317891-A.
PD 08-APR-1998.
PE 01-OCT-1997; 020890.
PR 14-AUG-1997; US-915503.
PR 01-OCT-1996; US-724643.
PR 18-APR-1997; US-844419.
PR 25-APR-1997; US-846017.
PR 06-MAY-1997; US-851843.
PR 09-MAY-1997; US-854050.
PR 14-AUG-1997; US-911312.
PR 14-AUG-1997; US-912951.
PA (GERO-) GERON CORP.
PA (UYTE-) UNIV TECHNOLOGY CORP.
PI Andrews WH, Czech TR, Chapman KB, Harley C, Lingner J,
PI Morin GB, Nakamura T, Harley CB;
DR WPI: 98-171633/16.
PT Pure and recombinant human Telomerase Reverse Transcriptase and its
PT variants - are useful in the diagnosis, prognosis and treatment of
PT cell proliferation conditions especially cancer and ageing
PS Example 1; Page 190; 387pp; English.
CC The present sequence represents a competitor DNA oligonucleotide
CC used in an example from the present invention which describes
CC human telomerase reverse transcriptase (hTERT). The present
CC invention also describes the following methods: (A) determining
CC whether a test compound is a modulator of hTERT, by detecting the change
CC in hTERT recombinant protein or polynucleotide, on administration of the
CC compound; (B) preparation of recombinant telomerase by contacting a
CC protein preparation of hTERT with a telomerase RNA component; (C)
CC detection of the hTERT RNA or protein in a sample by binding a relevant
CC probe to the sample and detecting the complex formed or in the case of
CC RNA detection, amplifying the product and correlating the presence of
CC complex or amplification product with presence of hTERT in the sample;
CC and (D) increasing the proliferation of a vertebrate cell by increasing
CC hTERT expression; and (E) the use of an agent that causes an increase in
CC cell vertebrate cell proliferation to create a medicament that inhibits
CC ageing. A protein preparation of hTERT and the polynucleotide encoding
CC hTERT can be used in the manufacture of medicaments for inhibiting the
CC effect of ageing or cancer. Inhibitors of telomerase activity can be
CC used to treat conditions that are associated with high telomerase
CC activity. A protein preparation of hTERT can also be used in the new
CC methods.
SQ Sequence 30 BP; 6 A; 4 C; 9 G; 11 T;

Query Match 67.8%; Score 12.2; DB 1; Length 30;
Best Local Similarity 82.4%; Pred. No. 7.3e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ggcacaaacagccta 17
||| ||| ||| ||| |||
DB 26 GCCACAAATCCACCTA 10

RESULT 10
Q12661/c
ID Q12661 standard; DNA; 48 BP.
AC Q12661:
DT 01-OCT-1991 (first entry)
DE SUC2 signal peptide.
KW sucrose invertase; yeast; *Pichia pastoris*; mutant MP-36; ds.
OS Synthetic.
PN EP-438200-A.
PD 24-JUL-1991.
PE 16-JAN-1991; 200074.
PR 16-JAN-1990; CU-000007.
PR 13-JUN-1990; CU-000104.
PR 12-JUL-1990; CU-000132.
PR 15-AUG-1990; CU-000142.
PA (INGE-) CENT INGEN GENETICA.
PI Herrera-Martinez LS, Yonng-Gonzalez V, Margolles-Clark E,
PI Delgado-Boada JM, Morales-Grillo J, Torres-Madrado IC,

PI Silva-Rodriguez A, Palfer-Reyes E;
DR WPI: 91-216886/30.
PT Expression of heterologous genes in *Pichia pastoris* yeast - by
PT using expression vectors and transformed microorganisms
PS Example 2; Fig 5; 24pp; English.
CC The 5' end overhangs the 3' end of the complementary strand by 4
CC bases; the 5' end of the complementary strand overhangs the 3' end
CC of the sense strand by the four bases 5'CATG 3'. The cohesive ends
CC are HindIII and NcoI sites, respectively. The synthetic sequence is
CC cloned into a yeast expression vector to give plasmid pps-7 which
CC contains the AOX1 promoter, the SUC2 signal sequence, a fragment of
CC the *S.cerevisiae* GAP operator (GAP), the *S.cerevisiae* HIS3 gene
CC and the 3' terminal fragment of the AOX gene for integration in the
CC yeast. Insertion of a heterologous gene occurs between the SUC2
CC signal sequence and GAP using the regenerated NcoI site.
CC See also Q12662.
SQ Sequence 48 BP; 8 A; 12 C; 10 G; 18 T;

Query Match 67.8%; Score 12.2; DB 1; Length 48;
Best Local Similarity 82.4%; Pred. No. 7.7e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ggcacaaacagccta 17
||| ||| ||| ||| |||
DB 32 GCTGCAAAACGACCA 16

RESULT 11
Q34057
ID Q34057 standard; DNA; 55 BP.
AC Q34057:
DT 02-FEB-1993 (first entry)
DE Downstream sequence of microsatellite from clone TGLM44.
KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
KM genetic mapping; traits; amplification; ss.
OS *Bos taurus*.
PN WO9213102-A.
PD 06-AUG-1992.
PR 15-JAN-1992; U00340.
PR 15-JAN-1991; US-642342.
PA (GENM-) GENMARK.
PI Georges M, Massey JM;
DR WPI: 92-284684/34.
PT Polymorphic bovine DNA markers - used in genetic identification,
PT gene mapping, and selective breeding
PS Table 7; Page 353; 517pp; English.
CC The sequence is that downstream of a bovine microsatellite sequence
CC obt'd. by screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (T6)n >9 microsatellites in the bovine genome is estimated at >100.
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellite (using the program
CC OPTIPRIM). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved in the determination of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 55 BP; 19 A; 16 C; 11 G; 9 T;

Query Match 67.8%; Score 12.2; DB 1; Length 55;
Best Local Similarity 82.4%; Pred. No. 7.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ggcacaaacagccta 17
||| ||| ||| ||| |||
DB 37 GCGACTAAACGACATA 53

RESULT 12

004915/c
ID 004915 standard; DNA: 57 BP.
AC 004915;
DT 23-OCT-1990 (first entry)
DE Wild type signal sequence of plasmid pING156.
KW Thaumatin; sweetener; sugar; sucrose; ds.
OS Thaumatooccus daniellii.
PN W09005775-A.
PD 31-MAY-1990.
PF 6-NOV-1989; 005018.
PR 8-NOV-1988; US-268702.
PS 4-SEP-1989; US-407416.
PT (ITGE-) Int Genetics Eng Inc.
PI Blair LC, Koduri RK, Lee JH, Weickmann JL;
DR WPI: 90-193426/25.
P-PSDB: R05497.
PT DNA encoding thaumatin 1 analogues with improved taste -
PT for use as sweeteners and tools for investigating the taste
response.
PS Example 5; Table 5; 65pp; English.
CC Thaumatin is 5000 times sweeter in taste than sucrose, and has
potential as sugar substitute. The plasmid contains promoter,
secretion signal and terminator sequences, and may be used in an
E.coli or S.cerevisiae expression system.
CC The modified signal sequence increases yield and decreases lag
time for the system.
SQ Sequence 57 BP; 9 A; 14 C; 10 G; 24 T;

Query Match 67.8%; Score 12.2; DB 1; Length 57;
Best Local Similarity 82.4%; Pred. No. 7.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ggcacaaaccagccta 17
||| ||||| ||||| |||
DB 44 GCTGCAAAACGACCAA 28

RESULT 13

004916/c
ID 004916 standard; DNA: 57 BP.
AC 004916;
DT 23-OCT-1990 (first entry)
DE Modified signal sequence of plasmid pING447.
KW Thaumatin; sweetener; sugar; sucrose; ds.
OS Thaumatooccus daniellii.
PN W09005775-A.
PD 31-MAY-1990.
PF 6-NOV-1989; 005018.
PR 8-NOV-1988; US-268702.
PS 4-SEP-1989; US-407416.
PT (ITGE-) Int Genetics Eng Inc.
PI Blair LC, Koduri RK, Lee JH, Weickmann JL;
DR WPI: 90-193426/25.
P-PSDB: R05498.
PT DNA encoding thaumatin 1 analogues with improved taste -
PT for use as sweeteners and tools for investigating the taste
response.
PS Example 5; Table 5; 65pp; English.
CC Thaumatin is 5000 times sweeter in taste than sucrose, and has
potential as sugar substitute. The plasmid contains promoter,
secretion signal and terminator sequences, and may be used in an
E.coli or S.cerevisiae expression system.
CC The modified signal sequence increases yield and decreases lag
time for the system.
SQ Sequence 57 BP; 11 A; 13 C; 11 G; 22 T;

Query Match 67.8%; Score 12.2; DB 1; Length 57;
Best Local Similarity 82.4%; Pred. No. 7.8e+02;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ggcacaaaccagccta 17
||| ||||| ||||| |||
DB 44 GCTGCAAAACGACCAA 28

RESULT 14

004917/c
ID 004917 standard; DNA: 57 BP.
AC 004917;
DT 23-OCT-1990 (first entry)
DE Modified signal sequence of plasmid pING448.
KW Thaumatin; sweetener; sugar; sucrose; ds.
OS Thaumatooccus daniellii.
PN W09005775-A.
PD 31-MAY-1990.
PF 6-NOV-1989; 005018.
PR 8-NOV-1988; US-268702.
PS 4-SEP-1989; US-407416.
PT (ITGE-) Int Genetics Eng Inc.
PI Blair LC, Koduri RK, Lee JH, Weickmann JL;
DR WPI: 90-193426/25.
P-PSDB: R05499.
PT DNA encoding thaumatin 1 analogues with improved taste -
PT for use as sweeteners and tools for investigating the taste
response.
PS Example 5; Table 5; 65pp; English.
CC Thaumatin is 5000 times sweeter in taste than sucrose, and has
potential as sugar substitute. The plasmid contains promoter,
secretion signal and terminator sequences, and may be used in an
E.coli or S.cerevisiae expression system.
CC The modified signal sequence increases yield and decreases lag
time for the system.
SQ Sequence 57 BP; 11 A; 13 C; 11 G; 22 T;

Query Match 67.8%; Score 12.2; DB 1; Length 57;
Best Local Similarity 82.4%; Pred. No. 7.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ggcacaaaccagccta 17
||| ||||| ||||| |||
DB 44 GCTGCAAAACGACCAA 28

RESULT 15

004918/c
ID 004918 standard; DNA: 57 BP.
AC 004918;
DT 23-OCT-1990 (first entry)
DE Modified signal sequence of plasmid pING449.
KW Thaumatin; sweetener; sugar; sucrose; ds.
OS Thaumatooccus daniellii.
PN W09005775-A.
PD 31-MAY-1990.
PF 6-NOV-1989; 005018.
PR 8-NOV-1988; US-268702.
PS 4-SEP-1989; US-407416.
PT (ITGE-) Int Genetics Eng Inc.
PI Blair LC, Koduri RK, Lee JH, Weickmann JL;
DR WPI: 90-193426/25.
P-PSDB: R05500.
PT DNA encoding thaumatin 1 analogues with improved taste -
PT for use as sweeteners and tools for investigating the taste
response.
PS Example 5; Table 5; 65pp; English.
CC Thaumatin is 5000 times sweeter in taste than sucrose, and has
potential as sugar substitute. The plasmid contains promoter,
secretion signal and terminator sequences, and may be used in an
E.coli or S.cerevisiae expression system.
CC The modified signal sequence increases yield and decreases lag
time for the system.

Sequence 57 BP; 10 A; 12 C; 11 G; 24 T;

Query Match 67.8%; Score 12.2; DB 1; Length 57;

Best Local Similarity 82.4%; Pred. No. 7.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 gcgacaaaccagccta 17
|| |||||
Db 44 GCTGCATAACGACCAA 28

Search completed: June 4, 2000, 16:24:10
Job time: 28906 sec


```
|||||
Db      3 AACGTAGTGTGATGATGATG 27

RESULT 6
ID      V84716
V84716  V84716 standard; DNA; 75 BP.
AC      V84716:
DE      27-APR-1999 (first entry)
KW      Monoclonal antibody /7F11 PCR primer P.
KW      Fab phase display library; multivalent polypeptide display library; detection;
KW      tumour-specific antigen; affinity screening; PCR primer; ss.
OS      Synthetic.
PN      MO9847343-A2.
PD      29-OCT-1998.
PR      03-APR-1998; 006704.
PR      04-APR-1997; US-835159.
PR      04-APR-1997; US-832985.
PA      (BIOS-) BIOSITE DIAGNOSTICS INC.
PI      Buehler J, Gray J, Valkiits G;
DR      WPI; 99-024016/02.
PT      Multivalent polypeptide display library production by selecting
PT      phages expressing two or more tags - particularly for Fab selection,
PT      producing specific, polyclonal populations and allowing DNA transfer
PT      without clonal isolation
PS      Example 22; Page 79; 104pp; English.
CC      This invention describes the production of a multivalent polypeptide
CC      display library. The described method is especially used to produce and
CC      screen Fab phase display libraries, particularly to isolate Fab with
CC      specific affinity for use as therapeutic or diagnostic agents, e.g. for
CC      detection of viral antigens or for delivering toxic substances to cells
CC      expressing tumour-specific antigens. Cell libraries (L2) are used to
CC      produce polypeptide libraries useful in diagnostic kits. The method
CC      enriches for members expressing multiple copies before affinity
CC      screening, i.e. it generates a diverse population of members expressing
CC      full-length coding sequences that may differ in binding affinity and
CC      specificity. Since the selected sublibrary contain a very high proportion
CC      of members with the required binding specificity, subcloning of DNA
CC      inserts from display to expression vectors can be done without clonal
CC      isolation. Since the population of Fab selected is polyclonal, different
CC      members may be reactive with different variants (e.g. escape mutants) of
CC      the target antigen. V84701-V84718 represents PCR primers used in the
CC      method of the invention.
SQ      Sequence 75 BP; 23 A; 12 C; 18 G; 22 T;

Query Match      55.1%; Score 24.8; DB 1; Length 75;
Best Local Similarity 92.9%; Pred. No. 1.1;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      8 ttagtgatggtgatgcatgaagaacc 35
        |||||||
Db      45 TTAGTGATGGTGGATGGATGACATCC 72

RESULT 7
ID      V93934/C
V93934/C V93934 standard; DNA; 69 BP.
AC      V93934:
DE      25-MAR-1999 (first entry)
DE      Oligonucleotide used to construct plasmid pMP151.
KW      Thrombopoietin protein; TPO; growth factor; early bone marrow stem cell;
KW      megakaryocyte precursor; megakaryocytic differentiation; chemotherapy;
KW      increased platelet count; thrombocytopenia; radiation treatment;
KW      liver cirrhosis; bone marrow hypoplasia; iron deficiency;
KW      disseminated intravascular coagulation; Flt-3; KL; CD34+ cell increase;
KW      myeloproliferative thrombocytotic disease;
KW      inflammation-related thrombocytosis; ss.
OS      Synthetic.
PN      MO9852598-A1.
PD      26-NOV-1998.
PR      28-JAN-1998; US-015016.
PR      21-MAY-1997; US-859767.
PA      (GETH ) GENENTECH INC.
PI      Cohen RL, Eaton DL, Jones AJ5, Jones DV, Powell MF,
PI      Sweeney TD, Thomas GR, Wagemaker G;
DR      WPI; 99-070128/06.
PT      Treating patients with, or at risk of, thrombocytopenia - by
PT      administering thrombopoietin in one or a few doses per day,
PT      particularly for subjects receiving radiation treatment or
PT      chemotherapy, speeds recovery and reduces need for platelet
PT      transfusions
PS      Example 3; Page 39; 105pp; English.
CC      The duplex formed by oligonucleotides V93934-95 was used to construct
CC      plasmid pMP151. This plasmid is a designed to express the first 155 amino
CC      acids of thrombopoietin protein (TPO) downstream of a 7 amino acid leader
CC      sequence, a 8 His tag and a factor Xa cleavage site. TPO is a growth
CC      factor acting directly on early bone marrow stem cells and megakaryocyte
CC      precursors (contrast granulocyte and/or macrophage colony stimulating
CC      factors which act at a later stage), resulting in megakaryocytic
CC      differentiation of bone marrow stem cells and increased platelet count.
CC      TPO can be used to treat a mammal with, or at risk of developing,
CC      thrombocytopenia. The treatment is particularly used for patients
CC      undergoing one or more cycles of radiation treatment and/or chemotherapy,
CC      but also those at risk from thrombocytopenia for other reasons such as
CC      liver cirrhosis, treatment to mobilise progenitor cells or accidental
CC      exposure to radiation. It may also be used to treat bone marrow
CC      hypoplasia, disseminated intravascular coagulation, any form of
CC      thrombocytopenia, possibly also myeloproliferative thrombocytotic
CC      diseases, inflammation-related thrombocytosis and iron deficiency.
CC      TPO, in conjunction with the cytokines Flt-3 and/or KL, can also be
CC      used to increase the number of CD34+ cells in a mammalian blood sample.
SQ      Sequence 69 BP; 25 A; 19 C; 9 G; 16 T;
```

```
PR      28-JAN-1998; US-015016.
PR      21-MAY-1997; US-859767.
PA      (GETH ) GENENTECH INC.
PI      Cohen RL, Eaton DL, Jones AJ5, Jones DV, Powell MF,
PI      Sweeney TD, Thomas GR, Wagemaker G;
DR      WPI; 99-070128/06.
PT      Treating patients with, or at risk of, thrombocytopenia - by
PT      administering thrombopoietin in one or a few doses per day,
PT      particularly for subjects receiving radiation treatment or
PT      chemotherapy, speeds recovery and reduces need for platelet
PT      transfusions
PS      Example 3; Page 39; 105pp; English.
CC      The duplex formed by oligonucleotides V93934-95 was used to construct
CC      plasmid pMP151. This plasmid is a designed to express the first 155 amino
CC      acids of thrombopoietin protein (TPO) downstream of a 7 amino acid leader
CC      sequence, a 8 His tag and a factor Xa cleavage site. TPO is a growth
CC      factor acting directly on early bone marrow stem cells and megakaryocyte
CC      precursors (contrast granulocyte and/or macrophage colony stimulating
CC      factors which act at a later stage), resulting in megakaryocytic
CC      differentiation of bone marrow stem cells and increased platelet count.
CC      TPO can be used to treat a mammal with, or at risk of developing,
CC      thrombocytopenia. The treatment is particularly used for patients
CC      undergoing one or more cycles of radiation treatment and/or chemotherapy,
CC      but also those at risk from thrombocytopenia for other reasons such as
CC      liver cirrhosis, treatment to mobilise progenitor cells or accidental
CC      exposure to radiation. It may also be used to treat bone marrow
CC      hypoplasia, disseminated intravascular coagulation, any form of
CC      thrombocytopenia, possibly also myeloproliferative thrombocytotic
CC      diseases, inflammation-related thrombocytosis and iron deficiency.
CC      TPO, in conjunction with the cytokines Flt-3 and/or KL, can also be
CC      used to increase the number of CD34+ cells in a mammalian blood sample.
SQ      Sequence 69 BP; 25 A; 19 C; 9 G; 16 T;

Query Match      54.7%; Score 24.6; DB 1; Length 69;
Best Local Similarity 87.1%; Pred. No. 1.3;
Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      9 tagtgatggtgatgcatgaagaaccaat 39
        |||||||
Db      49 TGTGTGATGGTGTGATGTAATGCCATAT 19

RESULT 8
ID      V93936/C
V93936/C V93936 standard; DNA; 69 BP.
AC      V93936:
DE      25-MAR-1999 (first entry)
DE      Oligonucleotide used to construct plasmid pMP202.
KW      Thrombopoietin protein; TPO; growth factor; early bone marrow stem cell;
KW      megakaryocyte precursor; megakaryocytic differentiation; chemotherapy;
KW      increased platelet count; thrombocytopenia; radiation treatment;
KW      liver cirrhosis; bone marrow hypoplasia; iron deficiency;
KW      disseminated intravascular coagulation; Flt-3; KL; CD34+ cell increase;
KW      myeloproliferative thrombocytotic disease;
KW      inflammation-related thrombocytosis; ss.
OS      Synthetic.
PN      MO9852598-A1.
PD      26-NOV-1998.
PR      28-JAN-1998; US-015016.
PR      21-MAY-1997; US-859767.
PA      (GETH ) GENENTECH INC.
PI      Cohen RL, Eaton DL, Jones AJ5, Jones DV, Powell MF,
PI      Sweeney TD, Thomas GR, Wagemaker G;
DR      WPI; 99-070128/06.
PT      Treating patients with, or at risk of, thrombocytopenia - by
PT      administering thrombopoietin in one or a few doses per day,
PT      particularly for subjects receiving radiation treatment or
PT      chemotherapy, speeds recovery and reduces need for platelet
PT      transfusions
PS      Example 3; Page 39; 105pp; English.
CC      The duplex formed by oligonucleotides V93936-97 was used to construct
```


Query Match 51.1%; Score 23; DB 1; Length 51;
 Best Local Similarity 83.9%; Pred. No. 4.5;
 Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 8 ttagtgatgtgtatgtatgaaagccaa 38
 |||
 DB 48 TTCGTGATGTGATGTGATGAATTCGAA 18

RESULT 15

VS1820
 ID VS1820 standard; DNA; 58 BP.
 AC VS1820;
 DT 04-MAR-1999 (first entry)
 DE Granulocytic Ehrlichia protein clone E46 PCR reverse primer.
 KM Granulocytic ehrlichiosis; Ehrlichia sp.; GE protein; infection; tick;
 KM diagnosis; vaccine; antigenic protein; antibody; immune response;
 KM PCR primer; ss.
 OS Synthetic.
 OS Ehrlichia sp.
 PN W09849313-A2.
 PD 05-NOV-1998.
 PE 24-APR-1998; U08265.
 PR 25-APR-1997; US-044933.
 PA (AQU1-) AQUILA BIOPHARMACEUTICALS INC.
 PI Beltz GA, Coughlin RT, Murphy CA, Storey J;
 DR WPI; 99-009432/01.
 PT New nucleic acid from the human granulocytic ehrlichiosis agent -
 PT and related antigenic proteins; vectors, transformed cells and
 PT antibodies; useful for diagnosis and in protective vaccines
 PS Example 15; Page 71; 154pp; English.
 CC The present sequence represents a PCR primer for granulocytic ehrlichia
 CC (GE) proteins used in an example from the present invention.
 CC GE nucleic acids, vectors and host cells are used for the recombinant
 CC production of GE proteins, and also in research to further characterise
 CC the proteins. GE protein-encoding nucleic acid molecules are detected
 CC by hybridisation to GE nucleic acid fragments or by using the fragments
 CC as primers for polymerase chain reaction (PCR) amplification. GE
 CC proteins, their immunogenic fragments, and GE nucleic acid molecules
 CC encoding them are used to generate an immune response against GE,
 CC specifically as (genetic) vaccines, especially to control ehrlichiosis
 CC in humans and dogs, but also to raise Ab and to study DNA-protein
 CC interactions. Ab are used to detect GE proteins by forming an immune
 CC complex in standard assays, and correspondingly GE proteins can detect
 CC specific antibodies, especially for diagnosis, assessment and prognosis
 CC of GE infection, or of contamination of biological samples with GE.
 SQ Sequence 58 BP; 15 A; 15 C; 17 G; 11 T;

Query Match 50.2%; Score 22.6; DB 1; Length 58;
 Best Local Similarity 86.2%; Pred. No. 6.4;
 Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 9 tagtgatgtgtatgtatgaaagccaa 37
 |||
 DB 12 TAGTGTGTGTGTGTGTGAAGCAAA 40

Search completed: June 4, 2000, 16:24:18
 Job time: 28914 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 08:20:48 ; Search time 1236.38 Seconds
(without alignments)
-13.376 Million cell updates/sec

Title: US-09-164-714-2

Perfect score: 17

Sequence: 1 gaycncarcntaygt 17

Scoring table: IDENTITY_NUC

Searched: Gapop 10.0 ; Gapext 1.0

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Database : Listing first 45 summaries

1: GenBml:*
2: gb_bal:*
3: gb_bal:*
4: gb_bal:*
5: gb_bal:*
6: gb_bal:*
7: gb_bal:*
8: gb_bal:*
9: gb_bal:*
10: gb_bal:*
11: gb_bal:*
12: gb_bal:*
13: gb_bal:*
14: gb_bal:*
15: gb_bal:*
16: gb_bal:*
17: gb_bal:*
18: gb_bal:*
19: gb_bal:*
20: gb_bal:*
21: gb_bal:*
22: gb_bal:*
23: gb_bal:*
24: gb_bal:*
25: gb_bal:*
26: gb_bal:*
27: gb_bal:*
28: gb_bal:*
29: gb_bal:*
30: gb_bal:*
31: gb_bal:*
32: gb_bal:*
33: gb_bal:*
34: gb_bal:*
35: gb_bal:*
36: gb_bal:*
37: gb_bal:*
38: gb_bal:*
39: gb_bal:*
40: gb_bal:*
41: gb_bal:*
42: gb_bal:*
43: gb_bal:*
44: gb_bal:*

45: gb_bal:*
46: gb_bal:*
47: gb_bal:*
48: gb_bal:*
49: gb_bal:*
50: gb_bal:*
51: gb_bal:*
52: gb_bal:*
53: gb_bal:*
54: gb_bal:*
55: gb_bal:*
56: gb_bal:*
57: gb_bal:*
58: gb_bal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
C	1	11.2	65.9	30	34	DRORN	K01289 D.melanogata	
C	2	11.2	65.9	52	14	SYNANVA4	M60112 Avian neov	
C	3	11.2	65.9	76	16	AB003918	AB003918 Hepatitis	
C	4	11.2	65.9	90	12	RATVPNP	K02434 Rat vasopre	
C	5	11.2	65.9	90	12	RATVPNP	K01495 Rat prepro	
C	6	11.1	64.7	21	5	AR030742	AR030742 Sequence	
C	7	11.1	64.7	21	5	AR030743	AR030743 Sequence	
C	8	11.1	64.7	75	5	I25131	I25131 Sequence	
C	9	11.1	64.7	90	5	I25143	I25143 Sequence	
C	10	10.6	62.4	20	5	AR012271	AR012271 Sequence	
C	11	10.6	62.4	20	5	I14969	I14969 Sequence	
C	12	10.6	62.4	20	5	I73689	I73689 Sequence	
C	13	10.6	62.4	35	5	I19619	I19619 Sequence	
C	14	10.6	62.4	71	5	AR017688	AR017688 Sequence	
C	15	10.2	60.0	21	5	A16125	A16125 Oligonucleo	
C	16	10.2	60.0	21	5	A16126	A16126 Oligonucleo	
C	17	10.2	60.0	21	5	I67007	I67007 Sequence	
C	18	10.2	60.0	21	5	I67008	I67008 Sequence	
C	19	10.2	60.0	30	5	A13425	A13425 Oligonucleo	
C	20	10.2	60.0	30	5	A13426	A13426 Oligonucleo	
C	21	10.2	60.0	30	5	A13427	A13427 Oligonucleo	
C	22	10.2	60.0	30	5	A13428	A13428 Oligonucleo	
C	23	10.2	60.0	30	5	A13541	A13541 Oligonucleo	
C	24	10.2	60.0	30	5	A13542	A13542 Oligonucleo	
C	25	10.2	60.0	30	5	A13543	A13543 Oligonucleo	
C	26	10.2	60.0	30	5	A13544	A13544 Oligonucleo	
C	27	10.2	60.0	30	5	A13545	A13545 Oligonucleo	
C	28	10.2	60.0	30	5	A13546	A13546 Oligonucleo	
C	29	10.2	60.0	30	5	A13547	A13547 Oligonucleo	
C	30	10.2	60.0	30	5	A13548	A13548 Oligonucleo	
C	31	10.2	60.0	30	5	A13568	A13568 Oligonucleo	
C	32	10.2	60.0	30	5	A13569	A13569 Oligonucleo	
C	33	10.2	60.0	30	5	I07186	I07186 Sequence	
C	34	10.2	60.0	30	5	I30350	I30350 Sequence	
C	35	10.2	60.0	32	5	A69250	A69250 Sequence	
C	36	10.2	60.0	32	5	A71927	A71927 Sequence	
C	37	10.2	60.0	32	5	E16812	E16812 PCR primer	
C	38	10.2	60.0	32	5	E16831	E16831 PCR primer	
C	39	10.2	60.0	35	5	A13549	A13549 Oligonucleo	
C	40	10.2	60.0	35	5	A13550	A13550 Oligonucleo	
C	41	10.2	60.0	36	5	I91771	I91771 Sequence	
C	42	10.2	60.0	47	7	MIRGGL3	X12479 Yeast mitoC	
C	43	10.2	60.0	60	5	A21412	A21412 Oligonucleo	
C	44	10.2	60.0	60	5	A21415	A21415 Oligonucleo	
C	45	10.2	60.0	60	5	A21420	A21420 Oligonucleo	

ALIGNMENTS

```

RESULT 1
DROORN/c      30 bp      RNA      INV      27-APR-1993
LOCUS         D.melanogaster 2S rRNA.
DEFINITION   K01289
ACCESSION    K01289.1 GI:174306
VERSION      2S ribosomal RNA; ribosomal RNA.
KEYWORDS     Drosophila melanogaster (adult and larval) rRNA.
SOURCE       Drosophila melanogaster
ORGANISM     Drosophila melanogaster
REFERENCE    1 (bases 1 to 30)
AUTHORS      Pavlakis,G.N., Jordan,B.R., Wurst,R.M. and Vournakis,J.N.
TITLE        Sequence and secondary structure of Drosophila melanogaster 5.8S
              and 2S rRNAs and of the processing site between them
JOURNAL      Nucleic Acids Res. 7, 2213-2238 (1979)
COMMENT      [1] also sequenced the 5.8S and 2S rRNA gene region, plus the 5.8S
              mature rRNA (see separate entries).
FEATURES     Location/Qualifiers
              source          1..30
                           /organism="Drosophila melanogaster"
                           /db_xref="taxon:7227"
              RNA            1..30
                           /note="2S rRNA"
BASE COUNT   6 a      3 c      10 g      11 t
ORIGIN       5' end of mature rRNA.

Query Match      65.9%; Score 11.2; DB 34; Length 30;
Best Local Similarity 62.5%; Pred. No. 1.4e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 gaygncarcntaygt 17
1:1111111111111111
Db 26 ACCCTCACCATATGTT 11

RESULT 2
SYNANVA4      52 bp      DNA      SYN      27-APR-1993
LOCUS         SYNANVA4
DEFINITION   Avian neovirus subgroup A DNA.
ACCESSION    M60112
VERSION      M60112.1 GI:207943
KEYWORDS     Synthetic viral DNA, clone QTE36-71.
SOURCE       Synthetic viral DNA, clone QTE36-71.
ORGANISM     Synthetic viral DNA, clone QTE36-71.
REFERENCE    1 (bases 1 to 52)
AUTHORS      Olsen,J.C., Bova-Hill,C., Grandgenett,D.P., Quinn,T.P.,
              Manfredi,J.P. and Swanson,R.
TITLE        Rearrangements in unintegrated retroviral DNA are complex and are
              the result of multiple genetic determinants
JOURNAL      J. Virol. 64, 5475-5484 (1990)
MEDLINE      91012792
FEATURES     Location/Qualifiers
              source          1..52
                           /organism="Synthetic Construct"
                           /db_xref="taxon:32630"
              LTR            28..52
                           /partial
BASE COUNT   15 a      15 c      9 g      13 t
ORIGIN

Query Match      65.9%; Score 11.2; DB 14; Length 52;
Best Local Similarity 62.5%; Pred. No. 1.4e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 gaygncarcntayg 16
11:1111111111111111

```

```

Db 19 GAGGCCCAATCTTANG 34

RESULT 3
AB003918/c    76 bp      RNA      VRL      10-DEC-1999
LOCUS         AB003918
DEFINITION   Hepatitis C virus genomic RNA for HVR1, partial cds.
ACCESSION    AB003918
VERSION      AB003918.1 GI:2116621
KEYWORDS     HVR1-19-1.
SOURCE       Hepatitis C virus cDNA to genomic RNA.
ORGANISM     Hepatitis C virus
REFERENCE    1 (sites)
AUTHORS      Yoshioka,K., Aiyama,T., Okumura,A., Takayanagi,M., Iwata,K.,
              Ishikawa,T., Nagai,Y. and Kakumu,S.
TITLE        Humoral immune response to the hypervariable region of hepatitis C
              virus differs between genotypes 1b and 2a
JOURNAL      J. Infect. Dis. 175 (3), 505-510 (1997)
MEDLINE      97193735
REFERENCE    2 (sites)
AUTHORS      Watanabe,K., Yoshioka,K., Ito,H., Ishigami,M., Takagi,K.,
              Utsunomiya,S., Kobayashi,M., Kishimoto,H., Yano,M. and Kakumu,S.
TITLE        The hypervariable region 1 protein of hepatitis C virus broadly
              reactive with sera of patients with chronic hepatitis C has a
              similar amino acid sequence with the consensus sequence
JOURNAL      Virology 264 (1), 153-158 (1999)
MEDLINE      20013225
REFERENCE    3 (bases 1 to 76)
AUTHORS      Yoshioka,K.
TITLE        Direct Submission
JOURNAL      Submitted (14-MAY-1997) to the DDBJ/EMBL/GenBank databases. Kentaro
              Yoshioka, Nagoya University School of Medicine, Third Department of
              Internal Medicine: 65 Tsuruma-cho, Showa-ku, Nagoya, Aichi
              466-0065, Japan (E-mail:kankenetsuru.med.nagoya-u.ac.jp,
              Tel:81-52-744-2190, Fax:81-52-744-2208)
FEATURES     Location/Qualifiers
              source          1..76
                           /organism="Hepatitis C virus"
                           /db_xref="taxon:11103"
                           <1..>76
                           /codon_start=2
                           /product="HVR1-19-1"
                           /protein_id="BAA20149.1"
                           /db_xref="GI:2116622"
                           /translation="HTHTVGGGSRARRLSFACLFYGP"
              CDS            13 a      26 c      22 g      15 t
BASE COUNT   13 a      26 c      22 g      15 t
ORIGIN

Query Match      65.9%; Score 11.2; DB 16; Length 76;
Best Local Similarity 62.5%; Pred. No. 1.4e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 gaygncarcntayg 16
11:1111111111111111
Db 23 GACCCCAACCGTATG 8

RESULT 4
RATVPN      90 bp      mRNA      ROD      27-APR-1993
LOCUS         RATVPN
DEFINITION   Rat vasopressin-neurophysin mRNA, partial 5' region.
ACCESSION    K02434
VERSION      K02434.1 GI:207675
KEYWORDS     neurophysin; vasopressin.
SOURCE       Rat (Brattleboro and Long-Evans) hypothalamus, cDNA to mRNA.
ORGANISM     Rattus norvegicus
REFERENCE    1 (bases 1 to 90)
AUTHORS      Majzoub,J.A., Pappey,A., Burg,R. and Habener,J.F.

```

TITLE Vasopressin gene is expressed at low levels in the hypothalamus of the Battlereboro rat

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81, 5296-5299 (1984)

MEDLINE 84296125

COMMENT Over the region sequenced (1-53), the Battlereboro gene appears to be identical to the Long-Evans gene [1].

FEATURES

source 1.90

sig_peptide /organism="Rattus norvegicus"

14..70 /db_xref="taxon:10116"

14..>90 /note="vasopressin-neurophysin signal pept"

CDS /note="vasopressin-neurophysin prepropeptide"

/codon_start=1

/protein_id="AA42343.1"

/db_xref="GI:207676"

/translation="MLNTTSLACFLSLALTSACYPQNC"

71..>90

mat_peptide /note="vasopressin"

13 a 36 c 16 g 25 t

BASE COUNT 13 a 36 c 16 g 25 t

ORIGIN unreported.

Query Match 65.9%; Score 11.2; DB 12; Length 90;

Best Local Similarity 62.5%; Pred. No. 1.4e+04;

Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gaycncarcntaygt 16

11:1111111111

Db 13 GATGCTCACTACACTAG 28

RESULT 5

LOCUS RATVPNP 90 bp mRNA ROD 27-APR-1993

DEFINITION Rat preprovasopressin-neurophysin gene, 5' end.

ACCESSION K01485

VERSION K01485.1 GI:207677

KEYWORDS neurophysin; vasopressin.

SOURCE rat (Long-Evans) adult female hypothalamus, cDNA to mRNA, clones cDNA-1 and cDNA-2.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 90)

AUTHORS Majzoub,J.A., Rich,A., van Boom,J.H. and Habener,J.F.

TITLE Vasopressin and oxytocin mRNA regulation in the rat assessed by hybridization with synthetic oligonucleotides

JOURNAL J. Biol. Chem. 258, 14061-14064 (1983)

MEDLINE 84061831

COMMENT When rats were given 2% saline solution to drink for 3 weeks the vasopressin levels in the hypothalamus increased 20-fold. There is one other entry under this title.

FEATURES

source 1..90

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

<1..>90

14..70 /note="vasopressin-neurophysin mRNA"

14..>90 /note="vasopressin-neurophysin signal peptide"

CDS /note="preprovasopressin-neurophysin"

/codon_start=1

/protein_id="AA42344.1"

/db_xref="GI:207678"

/translation="MLNTTSLACFLSLALTSACYPQNC"

71..>90

mat_peptide /note="vasopressin"

13 a 36 c 16 g 25 t

BASE COUNT 13 a 36 c 16 g 25 t

ORIGIN

Query Match 65.9%; Score 11.2; DB 12; Length 90;

Best Local Similarity 62.5%; Pred. No. 1.4e+04;

Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gaycncarcntaygt 16

11:1111111111

Db 13 GATGCTCACTACACTAG 28

RESULT 6

LOCUS AR030742/c 21 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 28 from patent US 5861298.

ACCESSION AR030742

VERSION AR030742.1 GI:5943956

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 21)

AUTHORS Adams,M.D., Blake,J.A., Debouck,C.M., Drake,F.H., Fitzgerald,L.M., Fraser,C.M., Gowen,M., Hastings,G.A., Kirkness,E.F., Lee,N.H. and Rood,J.

TITLE Cathepsin K gene

JOURNAL Patent: US 5861298-A 28 19-JAN-1999;

FEATURES

source 1..21

/organism="unknown"

BASE COUNT 5 a 5 c 5 g 5 t

ORIGIN

Query Match 64.7%; Score 11; DB 5; Length 21;

Best Local Similarity 64.7%; Pred. No. 1.9e+04;

Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 gaycncarcntaygt 17

11:1111111111

Db 18 GATGCTACCATATGCT 2

RESULT 7

LOCUS AR030743 21 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 29 from patent US 5861298.

ACCESSION AR030743

VERSION AR030743.1 GI:5943957

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 21)

AUTHORS Adams,M.D., Blake,J.A., Debouck,C.M., Drake,F.H., Fitzgerald,L.M., Fraser,C.M., Gowen,M., Hastings,G.A., Kirkness,E.F., Lee,N.H. and Rood,J.

TITLE Cathepsin K gene

JOURNAL Patent: US 5861298-A 29 19-JAN-1999;

FEATURES

source 1..21

/organism="unknown"

BASE COUNT 6 a 5 c 5 g 5 t

ORIGIN

Query Match 64.7%; Score 11; DB 5; Length 21;

Best Local Similarity 64.7%; Pred. No. 1.9e+04;

Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 gaycncarcntaygt 17

11:1111111111

Db 4 GATGCTACCATATGCT 20

RESULT 8
LOCUS 125131/c 75 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 21 from patent US 5547871.
ACCESSION 125131
VERSION 125131.1 GI:1605001
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 75)
AUTHORS Black,B.C. and Summers,M.D.
TITLE Heterologous signal sequences for secretion of insect controlling proteins
JOURNAL Patent: US 5547871-A 21 20-AUG-1996;
FEATURES Location/Qualifiers
source 1..75 /organism="unknown"
BASE COUNT 10 a 22 c 24 g 19 t
ORIGIN

Query Match 64.7%; Score 11; DB 5; Length 75;
Best Local Similarity 64.7%; Pred. No. 1.9e+04;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaygcncarcntaygt 17
Db 36 GAGGCCACGCCGACCT 20

RESULT 9
LOCUS 125143/c 90 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 39 from patent US 5547871.
ACCESSION 125143
VERSION 125143.1 GI:1605013
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 90)
AUTHORS Black,B.C. and Summers,M.D.
TITLE Heterologous signal sequences for secretion of insect controlling proteins
JOURNAL Patent: US 5547871-A 39 20-AUG-1996;
FEATURES Location/Qualifiers
source 1..90 /organism="unknown"
BASE COUNT 17 a 25 c 28 g 20 t
ORIGIN

Query Match 64.7%; Score 11; DB 5; Length 90;
Best Local Similarity 64.7%; Pred. No. 1.9e+04;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaygcncarcntaygt 17
Db 36 GAGGCCACGCCGACCT 20

RESULT 10
LOCUS AR012271 20 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 55 from patent US 5763254.
ACCESSION AR012271
VERSION AR012271.1 GI:3970261
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 20)
AUTHORS Woldike,H.Fabrizius, Hagen,F., Hjort,C.Malland and Hastrup,S.
TITLE Enzyme capable of degrading cellulose or hemicellulose
JOURNAL Patent: US 5686593-A 55 11-NOV-1997;
FEATURES Location/Qualifiers
source 1..20 /organism="unknown"
BASE COUNT 3 a 5 c 6 g 6 t
ORIGIN

Query Match 62.4%; Score 10.6; DB 5; Length 20;
Best Local Similarity 58.8%; Pred. No. 3.4e+04;

REFERENCE 1 (bases 1 to 20)
AUTHORS Woldike,H.Fabrizius, Hagen,F., Hjort,C.Malland and Hastrup,S.
TITLE Enzyme capable of degrading cellulose or hemicellulose
JOURNAL Patent: US 5763254-A 55 09-JUN-1998;
FEATURES Location/Qualifiers
source 1..20 /organism="unknown"
BASE COUNT 3 a 5 c 6 g 6 t
ORIGIN

Query Match 62.4%; Score 10.6; DB 5; Length 20;
Best Local Similarity 58.8%; Pred. No. 3.4e+04;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaygcncarcntaygt 17
Db 3 GATGCTCAGTCTACCT 19

RESULT 11
LOCUS 114969 20 bp DNA PAT 02-APR-1996
DEFINITION Sequence 55 from patent US 5457046.
ACCESSION 114969
VERSION 114969.1 GI:1249877
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 20)
AUTHORS Woldike,H.F., Hagen,F., Hjort,C. and Hastrup,S.
TITLE Enzyme capable of degrading cellulose or hemicellulose
JOURNAL Patent: US 5457046-A 55 10-OCT-1995;
FEATURES Location/Qualifiers
source 1..20 /organism="unknown"
BASE COUNT 3 a 5 c 6 g 6 t
ORIGIN

Query Match 62.4%; Score 10.6; DB 5; Length 20;
Best Local Similarity 58.8%; Pred. No. 3.4e+04;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaygcncarcntaygt 17
Db 3 GATGCTCAGTCTACCT 19

RESULT 12
LOCUS 173689 20 bp DNA PAT 03-APR-1998
DEFINITION Sequence 55 from patent US 5686593.
ACCESSION 173689
VERSION 173689.1 GI:3009830
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 20)
AUTHORS Woldike,H.Fabrizius, Hagen,F., Hjort,C.Malland and Hastrup,S.
TITLE Enzyme capable of degrading cellulose or hemicellulose
JOURNAL Patent: US 5686593-A 55 11-NOV-1997;
FEATURES Location/Qualifiers
source 1..20 /organism="unknown"
BASE COUNT 3 a 5 c 6 g 6 t
ORIGIN

Query Match 62.4%; Score 10.6; DB 5; Length 20;
Best Local Similarity 58.8%; Pred. No. 3.4e+04;

Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 gaygncarcntaygt 17
||:| | | | | | | | | |
Db 3 GATGCTCAGTGTACGT 19

RESULT 13
119619
LOCUS 119619 35 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 8 from patent US 5510099.
ACCESSION 119619
VERSION 119619.1 GI:1599974
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 35)
AUTHORS Short,J.M. and Kretz,P.L.
TITLE Mutagenesis testing using transgenic non-human animals carrying
test DNA sequences
JOURNAL Patent: US 5510099-A 8 23-APR-1996;
FEATURES Location/Qualifiers
source 1..35
BASE COUNT 9 a 11 c 8 g 7 t
ORIGIN

Query Match 62.4%; Score 10.6; DB 5; Length 35;
Best Local Similarity 58.8%; Pred. No. 3.4e+04;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 gaygncarcntaygt 17
||:| | | | | | | | | |
Db 10 GACGATCAGCCAAACGT 26

RESULT 14
AR017688 71 bp DNA PAT 05-DEC-1998
LOCUS AR017688
DEFINITION Sequence 89 from patent US 5780228.
ACCESSION AR017688
VERSION AR017688.1 GI:3973291
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 71)
AUTHORS Parma,D.H., Hicke,B., Bridonneau,P. and Gold,L.
TITLE High affinity nucleic acid ligands to lectins
JOURNAL Patent: US 5780228-A 89 14-JUL-1998;
FEATURES Location/Qualifiers
source 1..71
BASE COUNT 19 a 15 c 26 g 11 t
ORIGIN

Query Match 62.4%; Score 10.6; DB 5; Length 71;
Best Local Similarity 58.8%; Pred. No. 3.4e+04;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 gaygncarcntaygt 17
||:| | | | | | | | | |
Db 34 GACCCGCTACTTATGT 18

RESULT 15
A16125 21 bp DNA PAT 03-OCT-1994
LOCUS A16125/C
DEFINITION Oligonucleotide adaptor (SEQ ID NO: 8).
ACCESSION A16125

VERSION A16125.1 GI:640908
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 21)
AUTHORS Cornelissen,B.J.C., Melchers,L.S., Meulenhoff,E.J.S., Van
Roekel,J.S.C., Sela-Buurlage,M.B., Vloemans,A.A., Woloshuk,C.P.,
Bol,J.F. and Linthorst,H.J.M.
TITLE Fungal resistant plants, process for obtaining fungal resistant
plants and recombinant polynucleotides for use therein
JOURNAL Patent: EP 0440304-A 8 07-AUG-1991;
FEATURES MOGEN INTERNATIONAL N.V.; Rijksuniversiteit te Leiden
Location/Qualifiers
source 1..21
BASE COUNT 6 a 5 c 6 g 4 t
ORIGIN /organism="synthetic construct"
/db_xref="taxon:32630"

Query Match 60.0%; Score 10.2; DB 5; Length 21;
Best Local Similarity 60.0%; Pred. No. 6.2e+04;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 ygcncarcntaygt 17
:| | | | | | | | | |
Db 21 TGCACAGCCTCATGT 7

Search completed: June 4, 2000, 16:05:23
Job time: 27875 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 08:22:24 ; Search time 322.35 seconds
(without alignments)
13.195 Million cell updates/sec

Title: US-09-164-714-2

Perfect score: 17

Sequence: 1 gaycncarcntaygt 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	11	64.7	21	1 V09686	Human cathepsin K
C 2	11	64.7	21	1 V09687	Human cathepsin K
C 3	11	64.7	29	1 06652	PCR primer P1 to a
C 4	11	64.7	59	1 050214	c-fos position 272
C 5	11	64.7	63	1 080732	Drosophila melanog
C 6	11	64.7	63	1 V18214	Codon optimised si
C 7	11	64.7	75	1 067694	Codon-optimised D.
C 8	11	64.7	90	1 067706	Fragment B esteras
C 9	11	64.7	90	1 067329	Aair optimised seq
C 10	11	64.7	90	1 V18322	Oligomer for a cod
C 11	10.6	62.4	19	1 V72009	Electronic perturb
C 12	10.6	62.4	24	1 V11542	Recombinant MPO DN
C 13	10.6	62.4	26	1 V39432	Humanised anti-HM1
C 14	10.6	62.4	35	1 048146	Lambda 37225-37255
C 15	10.6	62.4	35	1 V39428	Humanised anti-HM1
C 16	10.6	62.4	38	1 V10832	Human MSH2 gene PC
C 17	10.6	62.4	38	1 V39427	Humanised anti-HM1
C 18	10.6	62.4	50	1 V39429	Humanised anti-HM1
C 19	10.6	62.4	50	1 V39430	Humanised anti-HM1
C 20	10.6	62.4	71	1 T57762	L-selectin family
C 21	10.6	62.4	75	1 011149	Probe GTR-1 based
C 22	10.2	60.0	18	1 T50625	Human CERP hairpin
C 23	10.2	60.0	20	1 X15745	PCR primer used to
C 24	10.2	60.0	30	1 N90827	Anti-R gene constr
C 25	10.2	60.0	30	1 N95064	HTLV-1 R homologue
C 26	10.2	60.0	30	1 068400	HTLV anti-R gene c
C 27	10.2	60.0	30	1 T47837	HTLV-1 R region ol
C 28	10.2	60.0	30	1 T72860	Alpha-1-antitryps
C 29	10.2	60.0	32	1 V49438	Primer AB028 for B
C 30	10.2	60.0	32	1 V53672	Primer for T. pall
C 31	10.2	60.0	36	1 097842	Human E2 CDNA prim
C 32	10.2	60.0	36	1 097839	Human E1 CDNA prim
C 33	10.2	60.0	36	1 T41762	Ubiquitin-actinact
C 34	10.2	60.0	40	1 072625	Primer P4, for sim

C 35	10.2	60.0	40	1 072627	Primer P6, for sim
C 36	10.2	60.0	50	1 T69252	Flt-1 gene probe B
C 37	10.2	60.0	60	1 010918	Tag DNA polymerase
C 38	10.2	60.0	61	1 010919	Tag DNA polymerase
C 39	10.2	60.0	63	1 010917	Tag DNA polymerase
C 40	10.2	60.0	63	1 010916	Tag DNA polymerase
C 41	10.2	60.0	63	1 010925	Tag DNA polymerase
C 42	10.2	60.0	63	1 057919	Diversified HIV-1
C 43	10.2	60.0	67	1 V82651	Oligonucleotide us
C 44	10.2	60.0	92	1 X23496	Human neutrophil c
C 45	10.2	60.0	97	1 040278	Sequence of synthe

ALIGNMENTS

```
RESULT 1
ID V09686 standard; DNA; 21 BP.
AC V09686:
DF 20-JUL-1998 (first entry)
DE Human cathepsin K gene exon 5 reverse PCR primer 4R.
KW Cathepsin K; human; osteoporosis; periodontal disease;
KW Paget's disease; Gaucher's disease; Alzheimer's disease;
KW central nervous system inflammation; hyperparathyroidism;
KW bone degradation; dental implant degradation; metastasis; tumour;
KW diagnosis; therapy; marker; PCR; primer; ss.
OS Synthetic.
OS Homo sapiens.
PN EP-812916-A2.
PD 17-DEC-1997.
PF 19-MAY-1997; 303395.
PR 26-AUG-1996; US-026083.
PR 14-JUN-1996; US-019942.
PR 17-JUN-1996; US-020273.
PR 26-AUG-1996; WO-014026.
PA (HUMA-) HUMAN GENOME SCI INC.
PA (GENO-) INST GENOMIC RES.
PA (SMK-) SMITHKLINE BEECHAM CORP.
PI Adams MD, Blake JA, Debonck CM, Drake FH, Fitzgerald LM,
PI Fraser CM, Gowen M, Hastings GA, Kirkness EF, Lee NH,
PI Rood J;
DR WPI: 98-034977/04.
PT DNA encoding human cathepsin K - useful for diagnosing and treating
PT diseases associated with cathepsin K e.g. osteoporosis, bone
PT degradation, metastatic tumours, etc
PS Example 1; Page 50; 84bp; English.
CC Oligonucleotide 4R comprises a reverse PCR primer from exon 5
CC (see V09670) of the human cathepsin K gene (see V09660). PCR
CC primers (see V09679-90) to adjacent exons of the cathepsin K
CC gene were used in the amplification of human genomic DNA. DNA
CC sequencing of intron-exon boundaries allowed sequencing of the
CC cathepsin genomic DNA. DNA encoding human cathepsin K is useful
CC for the diagnosis and treatment of e.g. osteoporosis, periodontal
CC disease, Paget's disease, Gaucher's disease, CNS inflammation,
CC Alzheimer's disease, hyperparathyroidism, bone degradation,
CC metastatic tumours, and degradation of bone implants and
CC prostheses, especially dental implants.
SQ Sequence 21 BP; 5 A; 5 C; 5 G; 6 T;

Query Match 64.7%; Score 11; DB 1; Length 21;
Best Local Similarity 64.7%; Pred. No. 4.3e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaycncarcntaygt 17
   ||:|:| | | | | | | |
Db 18 GATCCTACCAATAGT 2

RESULT 2
V09687
ID V09687 standard; DNA; 21 BP.
```

AC V09687; (first entry)
 DE 20-JUN-1998 (first entry)
 DE Human cathepsin K gene exon 5, 6 forward PCR primer 5+6F.
 KW Cathepsin K; human; osteoporosis; periodontal disease;
 KW Paget's disease; Gaucher's disease; Alzheimer's disease;
 KW central nervous system inflammation; hyperparathyroidism;
 KW bone degradation; dental implant degradation; metastasis; tumour;
 KW diagnosis; therapy; marker; PCR; primer; ss.
 OS Synthetic.
 OS Homo sapiens.
 PN EP-812916-A2.
 PD 17-DEC-1997.
 PF 19-MAY-1997; 303395.
 PR 26-AUG-1996; US-026083.
 PR 14-JUN-1996; US-019942.
 PR 17-JUN-1996; US-020273.
 PR 26-AUG-1996; WO-0114026.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (GENO-) INST GENOMIC RES.
 PA (SMK-) SMITHKLINE BEECHAM CORP.
 PI Adams MD, Blake JA, Debouck CM, Drake FH, Fitzgerald LM,
 PI Fraser GM, Gowen M, Hastings GA, Kirkness EF, Lee NH,
 PI Rood J.
 DR WPI: 96-034977/04.
 PT DNA encoding human cathepsin K - useful for diagnosing and treating
 PT diseases associated with cathepsin K e.g. osteoporosis, bone
 PT degradation, metastatic tumours, etc
 PS Example 1: Page 51: 84pp: English.
 CS Oligonucleotide 5+6F comprises a forward PCR primer from exon 5
 CC (see V09670) of the human cathepsin K gene (see V09660). PCR
 CC primers (see V09679-90) to adjacent exons of the cathepsin K
 CC gene were used in the amplification of human genomic DNA. DNA
 CC sequencing of intron-exon boundaries allowed sequencing of the
 CC cathepsin genomic DNA. DNA encoding human cathepsin K is useful
 CC for the diagnosis and treatment of e.g. osteoporosis, periodontal
 CC disease, Paget's disease, Gaucher's disease, hyperparathyroidism,
 CC Alzheimer's disease, hyperparathyroidism, bone degradation,
 CC metastatic tumours, and degradation of bone implants and
 CC prostheses, especially dental implants.
 SQ Sequence 21 BP; 6 A; 5 C; 5 G; 5 T;

Query Match 64.7%; Score 11; DB 1; Length 21;
 Best Local Similarity 64.7%; Pred. No. 4.9e+02;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 gaycncarcctatgt 17
 ||:|| ||:|| ||:||
 DB 4 GAWGCTACCATATGT 20

RESULT 3
 ID 06652/c
 AC 06652; standard; DNA: 29 BP.
 DE 19-JAN-1995 (first entry)
 DE PCR primer P1 to amplify human dystrophin gene.
 KW Polymerase chain reaction amplification; human dystrophin gene; PCR;
 KW ss.
 OS Synthetic.
 PN SUI792972-A.
 PD 07-FEB-1993.
 PF 15-MAY-1990; 826213.
 PR 15-MAY-1990; SU-826213.
 PA (MOLE-) MOL DIAGNOSIS METHODS RES CENTRE.
 PA Evgrafov OV, Polyakov AV;
 DR WPI: 94-181019/22.
 PT Amplification of DNA fragment of unknown sequence - giving wider
 PT functional possibilities because of increased distance between
 PT DNA fragment of known sequence and multiplied DNA fragment of
 PT unknown sequence
 PS Example: Column 4; 4pp; Russian.
 CC Two oligonucleotide primers P1 and P2 (06652 and 06653,

CC respectively) were synthesised. The primers were used to amplify
 CC the 5'-terminal region of the human dystrophin gene, the sequence
 CC of which was unknown. Primer P1 was designed to hybridise to a
 CC region between two restriction sites, i.e. NcoI (site 749) and
 CC EcoRI (site 701). The target DNA is cut with one restriction enzyme
 CC and then circularised. Cutting the circular DNA with the second
 CC restriction enzyme results in linear DNA in which the region to be
 CC amplified is flanked by binding sites for P1 (sense strand) and P2
 CC (complementary strand).
 SQ Sequence 29 BP; 10 A; 5 C; 6 G; 8 T;

Query Match 64.7%; Score 11; DB 1; Length 29;
 Best Local Similarity 64.7%; Pred. No. 5.1e+02;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 gaycncarcctatgt 17
 ||:|| ||:|| ||:||
 DB 17 GAAGCAGACCTATAT 1

RESULT 4
 ID 050214 standard; DNA: 50 BP.
 AC 050214;
 DE 06-MAY-1994 (first entry)
 DE c-fos position 2721-2770 INS mutagenic oligonucleotide.
 KW HIV-1; p17gag; inhibition; gag; M1; M2; M3; M4; silent mutation; mRNA;
 KW vector; p17; point mutation; p17M1234; Htat cells; gene replacement;
 KW inhibitory/instability signal; INS; stability; utilisation; vaccine;
 KW interferon; interleukin; fos proto-oncogene protein; growth factor;
 KW env; attenuated; AIDS; AIDS-related disease; latent infection;
 KW gene therapy; Human immunodeficiency virus type 1; p24;
 KW long terminal repeat; LTR; ss.
 OS Synthetic.
 PN WO9320212-A.
 PD 14-OCT-1993.
 PF 29-MAR-1993; U02908.
 PR 27-MAR-1992; US-858747.
 PA (USSH-) US SEC DEPT HEALTH.
 PA Felber BK, Pavlakis GN;
 PI WPI: 93-336919/42.
 DR Eliminating inhibitory-instability regions in mRNA to improve
 PT stability and expression - by making multiple point mutations
 PT within A-T rich regions
 PS Example 3: Page 70; 117pp: English.
 CC The sequences given in 050202-70 are oligonucleotides which were
 CC used to illustrate the method of the invention for the mutagenesis of
 CC inhibitory/instability signals (INS). Mutation in an INS produced
 CC using oligonucleotides such as these, increases the stability and/or
 CC utilisation of mRNA without changing its protein coding capacity, or
 CC if the sequence is changed, its function is maintained. Other genes
 CC encoding such mRNA molecules include growth factor, interferon,
 CC interleukin, fos proto-oncogene protein and HIV env and gag gene
 CC proteins. Nucleic acid constructs in which INS function has been
 CC impeded, can be used as vaccines, esp. against AIDS and AIDS-related
 CC diseases by preventing HIV from establishing a latent infection, as
 CC is possible using the INS, and thus escaping immune system
 CC surveillance. The constructs may also be used in gene therapy for
 CC gene replacement by homologous recombination with a target gene in
 CC situ. See also 050200-02.
 SQ Sequence 50 BP; 16 A; 13 C; 13 G; 8 T;

Query Match 64.7%; Score 11; DB 1; Length 50;
 Best Local Similarity 64.7%; Pred. No. 5.5e+02;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 gaycncarcctatgt 17
 ||:|| ||:|| ||:|| ||:||
 DB 19 GAAGCAGACCTATACGT 35

RESULT 5
080732/c
ID 080732 standard; DNA; 63 BP.
AC 080732;
DT 22-AUG-1995 (first entry)
DE Drosophila melanogaster esterase-6 signal peptide gene.
KW Esterase-6 signal peptide gene; Drosophila melanogaster;
KM modular expression vectors; ss.
OS Drosophila melanogaster.
PN W09428114-A.
PD 08-DEC-1994.
PF 27-MAY-1994; U06079.
PR 28-MAY-1993; US-070164.
PA (AMCY) AMERICAN CYANAMID CO.
PI Dierks PM, Webb NR;
DR WPI; 95-022788/03.
PT Recombinant double stranded DNA insect virus - comprising a
PT modular expression vector and a direct ligation virus vector
PS Disclosure; Page 85; 114pp; English.
CC 080732 encodes the native esterase-6 signal peptide from Drosophila
CC melanogaster. This sequence was used in the construction of a modular
CC expression vector, which was combined with a direct ligation virus
CC vector to produce a recombinant ds DNA insect virus. The ability to
CC directly insert foreign DNA into the virus is advantageous, as no
CC transfer or plasmid vectors are required.
SQ Sequence 63 BP; 9 A; 14 C; 22 G; 18 T;

Query Match 64.7%; Score 11; DB 1; Length 63;
Best Local Similarity 64.7%; Pred. No. 5.7e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaycncarcantaygt 17
||| ||| ||| ||| |||
Db 24 GAGGCCCGAGCCGACGT 8

RESULT 6
V18214/c
ID V18214 standard; DNA; 63 BP.
AC V18214;
DT 02-SEP-1998 (first entry)
DE Codon optimised signal sequence of esterase-6.
KW Codon optimised; North African scorpion; Insect toxin; Aair;
KW Insect cell; expression; plant protection; Insect damage;
KW Insect virus; nuclear polyhedrosis virus; granulosis virus;
KW non-occluded virus; entomopox virus; ss.
OS Synthetic.
PN A09748372-A.
PD 12-MAR-1998.
PF 15-DEC-1997; 048372.
PR 25-JAN-1993; US-009264.
PA (AMCY) AMERICAN CYANAMID CO.
PI Black BC, Brennan LA, Dierks PM;
DR WPI; 98-241214/22.
PT Isolated, codon optimised nucleic acid sequence encoding Androctonus
PT australis insect toxin - useful for protecting plants from insect
PT damage by delivering insect virus containing it
PS Disclosure; Page 22; 81pp; English.
CC V18209-15 represent codon optimised signal sequences used in the
CC course of the invention. The specification describes a codon optimised
CC sequence for the Androctonus australis (North African scorpion) insect
CC toxin (Aair) gene. A. australis binds to sodium channels in insects and
CC causes contractile paralysis in insect larvae. Codon usage of the wild
CC type gene (see V18208) is optimised for expression in insect cells (see
CC V18207). The sequence is combined with a secretory signal sequence, also
CC optimised for expression in insect cells (see V18209-15). The optimised
CC sequence is useful for producing recombinant Aair, especially in insect
CC cells, or for protecting plants from insect damage. The toxin gene is
CC delivered by an insect virus, such as nuclear polyhedrosis virus,
CC granulosis virus, non-occluded virus, or entomopox virus.
SQ Sequence 63 BP; 9 A; 14 C; 22 G; 18 T;

Query Match 64.7%; Score 11; DB 1; Length 63;
Best Local Similarity 64.7%; Pred. No. 5.7e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaycncarcantaygt 17
||| ||| ||| ||| |||
Db 24 GAGGCCCGAGCCGACGT 8

RESULT 7
067694/c
ID 067694 standard; DNA; 75 BP.
AC 067694;
DT 13-MAR-1995 (first entry)
DE Codon-optimised D. melanogaster esterase-6 signal sequence DNA.
KW Codon optimised; native; B. mori; pMHPC-12; signal peptide; chorion;
KW Insect controlling protein; toxin; Aair; baculovirus; ACMPV; Cry IVD;
KW expression; secretion; toxin-induced paralysis; cuticle; apolipophorin;
KW sex-specific; adipokinetic; esterase-6; D. melanogaster; neuropeptide;
KW M. sexta; enzyme; pyrenoxen tritici; B. thuringiensis; diuretic hormone;
KW eclosion hormone; prothoracicotropic hormone; adipokinetic hormone;
KW proctolin; juvenile hormone esterase; ss.
OS Drosophila melanogaster.
FH Key location/Qualifiers
FT cds 13..75
FT /tag= a
FT /product= Esterase-6 signal sequence
PN EP-608696-A.
PD 03-AUG-1994.
PF 10-JAN-1994; 100265.
PR 25-JAN-1993; US-009265.
PA (AMCY) AMERICAN CYANAMID CO.
PI Black BC;
DR WPI; 94-242108/30.
DR P-PDB: R56466.
PT Heterologous signal sequences for secretion of insect controlling
PT proteins - useful to protect plants from insect pests
PS Disclosure; Page 38; 69pp; English.
CC The sequences given in 067694-95 represent the codon optimised and
CC native coding sequences for the D. melanogaster esterase-6 signal
CC peptide, respectively. This signal peptide sequence may be used with
CC a DNA sequence encoding an insect controlling protein, esp. the insect-
CC specific toxin Aair. The fusion sequence may be introduced into an
CC insect virus, such as the baculovirus ACMPV. The insertion of the
CC Aair gene and the heterologous signal sequence into a baculovirus
CC results in the expression and secretion of the toxin. A susceptible
CC insect which ingests such a modified baculovirus will cease feeding
CC on plants due to toxin-induced paralysis at an earlier time than an
CC insect which ingests a wild-type baculovirus, thus reducing crop
CC damage.
SQ Sequence 75 BP; 10 A; 22 C; 24 G; 19 T;

Query Match 64.7%; Score 11; DB 1; Length 75;
Best Local Similarity 64.7%; Pred. No. 5.8e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaycncarcantaygt 17
||| ||| ||| ||| |||
Db 36 GAGGCCCGAGCCGACGT 20

RESULT 8
067706/c
ID 067706 standard; DNA; 90 BP.
AC 067706;
DT 13-MAR-1995 (first entry)
DE Fragment B esterase-6 oligomer; encodes esterase-6 signal sequence.
KW Codon optimised; native; B. mori; pMHPC-12; signal peptide; chorion;
KW Insect controlling protein; toxin; Aair; baculovirus; ACMPV; Cry IVD;
KW expression; secretion; toxin-induced paralysis; cuticle; apolipophorin;
KW sex-specific; adipokinetic; esterase-6; D. melanogaster; neuropeptide;

KM M. sexta; enzyme; Pyemotes tritici; B. thuringiensis; diuretic hormone;
KM eclosion hormone; prothoracicotropic hormone; adipokinetic hormone;
KM proctolin; juvenile hormone esterase; ss.
OS Drosophila melanogaster.
PN EP-608696-A.
PD 03-AUG-1994.
PF 10-JAN-1994; 100265.
PR 25-JAN-1993; US-009265.
PA (AMCY) AMERICAN CYANAMID CO.
PI Black BC;
DR WPI: 94-242108/30.
PT Heterologous signal sequences for secretion of insect controlling
PT proteins - useful to protect plants from insect pests
PS Example 2; Page 14; 69pp; English.
CC The sequences given in Q67700-10 are fragments which were used in the
CC construction of eight heterologous signal sequence-codon optimised Aairt
CC toxin gene cassettes. The cassettes were constructed in two pieces, a
CC "B" fragment unique for each construct, consisting of DNA coding for one
CC of eight heterologous signal sequences, plus the amino terminal portion
CC of the toxin coding region, and an "A" fragment which is the same for
CC each construct and encodes the remainder of the toxin coding gene. Each
CC of the fragments "A" and "B" were synthesised by annealing a pair of
CC oligomers containing a 15 bp overlap to complete the double stranded
CC molecule. Fusion constructs such as these, may be introduced into an
CC insect virus, such as the baculovirus AcMNPV. The insertion of the Aairt
CC gene and the heterologous signal sequence into a baculovirus results in
CC the expression and secretion of the toxin. A susceptible insect which
CC ingests such a modified baculovirus will cease feeding on plants due to
CC toxin-induced paralysis at an earlier time than an insect which ingests
CC a wild-type baculovirus, thus reducing crop damage.
SQ Sequence 90 BP; 17 A; 25 C; 28 G; 20 T;

Query Match 64.7%; Score 11; DB 1; Length 90;
Best Local Similarity 64.7%; Pred. No. 5.9e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 gagycncarcntaygt 17
|||11:111111
DB 36 GAGGCCACGCCGACGT 20
RESULT 9
Q67329/c
ID Q67329 standard; DNA; 90 BP.
AC Q67329;
DE 11-APR-1995 (first entry)
KW Aairt optimised sequence; fragment B, esterase-6 signal sequence oligomer.
KM Codon optimisation; Androctonus australis; insect; toxin; Aairt;
KM contractile paralysis; virus; expression; crop plant; ss.
OS Synthetic.
PN AU9453967-A.
PD 28-JUL-1994.
PF 24-JAN-1994; 053967.
PR 25-JAN-1993; US-009264.
PA (AMCY) AMERICAN CYANAMID CO.
PI Black BC, Brennan LA, Dierks PM;
DR WPI: 94-272167/34.
PT Codon optimised DNA for Androctonus australis insect toxin - used
PT to transform hosts, partic. insect viruses for use in protecting
PT plants against insects
PS Example 2; Page 33; 82pp; English.
CC The sequences given in Q67323-33 represents oligomers which were
CC used in the construction of eight heterologous signal sequence-codon
CC optimised Aairt toxin gene cassettes. The cassettes were synthesised and
CC assembled in two pieces. A "B" fragment, unique in each construct,
CC consists of DNA coding for one of eight heterologous signal sequences
CC plus the amino terminal portion of the toxin, and an "A" fragment which
CC is the same for each construct and encodes the remainder of the toxin
CC coding sequence. Each of fragments A and B is made by annealing a pair
CC oligomers. The constructs contain the codon optimised sequence derived
CC from A. australis which encodes insect toxin (Aairt). The Aairt causes
CC contractile paralysis at the ng to mcg range in insect larvae.

CC Insect viruses contg. this codon optimised Aairt gene may be used to
CC kill insects to reduce damage to crops. The use of the codon
CC optimised sequence enhances expression levels of the Aairt by insect
CC viruses.
SQ Sequence 90 BP; 17 A; 25 C; 28 G; 20 T;

Query Match 64.7%; Score 11; DB 1; Length 90;
Best Local Similarity 64.7%; Pred. No. 5.9e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 gagycncarcntaygt 17
|||11:111111
DB 36 GAGGCCACGCCGACGT 20

RESULT 10
V18222/c
ID V18222 standard; DNA; 90 BP.
AC V18222;
DE 02-SEP-1998 (first entry)
KW Oligomer for a codon optimised signal sequence from esterase-6.
KW Codon optimised; North African scorpion; insect toxin; Aairt;
KW insect cell; expression; plant protection; insect damage;
KW insect virus; nuclear polyhedrosis virus; granulosis virus;
KW non-occluded virus; entomopox virus; ss.
OS Synthetic.
PN AU9748372-A.
PD 12-MAR-1998.
PF 15-DEC-1997; 048372.
PR 25-JAN-1993; US-009264.
PA (AMCY) AMERICAN CYANAMID CO.
PI Black BC, Brennan LA, Dierks PM;
DR WPI: 98-241214/22.
PT Isolated, codon optimised nucleic acid sequence encoding Androctonus
PT australis insect toxin - useful for protecting plants from insect
PT damage by delivering insect virus containing it
PS Example 2; Page 33; 81pp; English.
CC The present sequence represents an oligomer comprising a codon optimised
CC signal sequence from esterase-6, and the amino terminal portion of the
CC codon optimised sequence for the Androctonus australis (North African
CC scorpion) insect toxin (Aairt) gene. This fragment is combined with a
CC second fragment containing the remainder of the toxin coding region. The
CC optimised sequence is useful for producing recombinant Aairt, especially
CC in insect cells, or for protecting plants from insect damage. The toxin
CC gene is delivered by an insect virus, such as nuclear polyhedrosis virus,
CC granulosis virus, non-occluded virus, or entomopox virus.
SQ Sequence 90 BP; 17 A; 25 C; 28 G; 20 T;

Query Match 64.7%; Score 11; DB 1; Length 90;
Best Local Similarity 64.7%; Pred. No. 5.9e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 gagycncarcntaygt 17
|||11:111111
DB 36 GAGGCCACGCCGACGT 20

RESULT 11
V72009
ID V72009 standard; DNA; 19 BP.
AC V72009;
DE 29-MAR-1999 (first entry)
KW Electronic perturbation catalysis oligonucleotide probe HLA 378.
KW Electronic perturbation analysis; hybridisation analysis; match hybrid;
KW mismatch hybrid; detection; clinical assay; optoelectronic device;
KW optical memory; nanofabrication; synthesis; self-assembly; denaturation;
KW self-organising; fluorescence; denaturation; probe; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT modified base 19 /*tag= a

/note- "3'-end modified by presence of Biotin"

FT MO9851819-A1.
PN 19-NOV-1998.
PD 07-MAY-1998; 009357.
PR 14-MAY-1997; 05-855058.
PA (NANO-) NANOGEN INC.
PI Heller MJ, O Connell JP, Sosnowski RG, Tu E;
DR MPI: 99-059702/05.
PT Hybridisation analysis using electronic stringency control device
PT based on changes in fluorescence - when an electric field is applied
PT to the hybrid, used e.g. for sequencing or for detecting single
PT mismatch mutations, also electronic perturbation catalysis
PS Example 1: Page 28; 56pp: English.
CC This sequence represents a probe used in a method for studying
CC hybridisation analysis by detecting electronic denaturation using
CC electronic perturbation analysis. The method is particularly used for
CC sequencing and to discriminate between match and mismatch hybrids,
CC particularly to detect point mutations, deletions, inserts, repeat
CC regions, single nucleotide polymorphisms, translocations, intron/exon
CC junctions, etc. The same principle may be used for most molecular
CC biological processes, e.g. antibody-antigen reactions, cell typing and
CC separation, enzymatic and other clinical assays, also in optoelectronic
CC devices and optical memory materials. The method is applicable to any
CC type of reaction, e.g. nanofabrication or other self-assembling or
CC self-organising processes, synthesis of nucleic acids, peptides, polymers
CC etc. The method relies on the observation that perturbation of a
CC fluorescence signal (particularly a peak) occurs at the electric power
CC level which causes denaturation or dehybridisation. The method is very
CC rapid, e.g. match/mismatch hybrids can be distinguished in less than a
CC minute, e.g. 5 sec. It permits use of long probes (over 20 bases) and
CC probe specificity is determined by the position of the label. Since
CC analysis does not require removal of mismatched probes, the sample can be
CC analysed repeatedly to improve assay statistics.
SQ Sequence 19 BP; 4 A; 4 C; 7 G; 4 T;

Query Match 62.4%; Score 10.6; DB 1; Length 19;
Best Local Similarity 58.8%; Pred. No. 8.5e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaycncarccontaygt 17
||:| ||:| ||:| ||:|
DB 1 GATGACGAGCTTACGT 17

RESULT 12
V11542/c
ID V11542 standard; DNA; 24 BP.
AC V11542;
AT 30-JUL-1998 (first entry)
DE Recombinant MPO DNA H7b 3'-end primer.
KM Myeloperoxidase; MPO; MPO-ANCA; immunoreaction; IgG; diagnosis;
KM Myeloperoxidase-specific anti-neutrophil cytoplasmic antibody; primer;
KM rheumatoid arthritis; systemic lupus erythematosus; nephritis; ss.
OS Synthetic.
PN WO9807848-A1.
PD 26-FEB-1998.
PR 21-AUG-1997; J02910.
PR 21-AUG-1996; JP-219855.
PA (TEIKO) TEIKOKU HORMONE MFG CO LTD.
PI Suzuki K, Tanokura M;
DR MPI: 98-169154/15.
PT Recombinant myeloperoxidase (MPO) fragments reactive with MPO
PT auto-antibody - for use in diagnosis of auto-immune diseases by
PT auto-antibody assay in blood samples
PS Disclosure: Page 9; 45pp: Japanese.
CC Primers V11523-V11546 are used in a method which produces recombinant
CC myeloperoxidase (MPO) fragments that are used to assay MPO reactivity
CC with a myeloperoxidase-specific anti-neutrophil cytoplasmic antibody
CC (MPO-ANCA). The MPO fragments may be fused with a histidine hexamer,
CC with maltose-binding protein or with glutathione S-transferase (GST).
CC The MPO fragments may be used for assay of MPO-ANCA in biological
CC samples such as blood by an immunoreaction of the sample with immobilised

CC MPO fragment, then a further reaction with labelled anti-human IgG
CC antibody (e.g. using a horseradish peroxidase, alkali phosphatase or
CC beta-galactosidase label), then a final assay of the bound label (e.g.
CC by reaction with o-phenylenediamine for the peroxidase, or with
CC p-nitrophenyl phosphate for the phosphatase or galactosidase). This assay
CC of MPO-ANCA in biological samples can be used for the diagnosis of
CC autoimmune diseases such as rheumatoid arthritis, systemic lupus
CC erythematosus and autoimmune nephritis.
SQ Sequence 24 BP; 5 A; 5 C; 8 G; 6 T;

Query Match 62.4%; Score 10.6; DB 1; Length 24;
Best Local Similarity 58.8%; Pred. No. 8.8e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaycncarccontaygt 17
||:| ||:| ||:| ||:|
DB 17 GATGCCACGCTTACGT 1

RESULT 13
V39432
ID V39432 standard; DNA; 26 BP.
AC V39432;
AT 21-SEP-1998 (first entry)
DE Humanised anti-HM1.24 antibody PCR primer SEQ ID NO:93.
KW Mouse; human; humanised; anti-HM1.24 antibody; myeloma; FR; CDR;
KM framework region; complementarity determining region; antigenicity;
KM PCR primer; ss.
OS Synthetic.
OS Mus sp.
OS Homo sapiens.
PN WO9814580-A1.
PD 03-APR-1998.
PR 03-OCT-1997; J03553.
PR 04-OCT-1996; JP-264756.
PA (CHUS) CHUGAI SEIYAKU KK.
PI Koshihara Y, Kosaka M, Ohtomo T, Ono K, Tsuchiya M,
PI Yoshimura Y;
DR MPI: 98-286421/25.
PT Humanised anti-HM1.24 antibody - for treatment of myeloma
PS Example 9: Page 147; 210pp: Japanese.
CC A humanised anti-HM1.24 antibody has been developed which comprises
CC human L and H chain C regions, and L and/or H chain V regions
CC containing material originating in mouse anti-HM1.24 antibody. The V
CC regions contain framework (FR) regions of human origin and
CC complementarity determining regions (CDR) of mouse origin, leading to
CC a reshaped humanised antibody. The C regions are human Ck (L-chain) and
CC human C gamma (especially C gamma 1) (H-chain). The FR regions of the
CC L chain V region are derived from human subtype HSG1 (e.g. from human
CC antibody RE1) and the FR regions of the H chain V region are derived
CC from human subtype HSG1 (e.g. FR1-3 from human antibody HG3 and FR4
CC from human antibody JH6). The present sequence represents a PCR primer
CC used in an example from the present invention. The antibodies are used
CC for the treatment of myeloma, especially by injection, intravenously,
CC intramuscularly or subcutaneously. The antibodies are used at 0.01-1000
CC (especially 5-100) mg/kg body weight. The humanised antibody has low
CC antigenicity and is therefore effective therapeutically in humans.
SQ Sequence 26 BP; 6 A; 7 C; 6 G; 7 T;

Query Match 62.4%; Score 10.6; DB 1; Length 26;
Best Local Similarity 58.8%; Pred. No. 8.9e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaycncarccontaygt 17
||:| ||:| ||:| ||:|
DB 9 GATGCTGAGCTTACGT 25

RESULT 14
Q48146
ID Q48146 standard; DNA; 35 BP.

```

AC 048146:
DE 10-FEB-1994 (first entry)
DE Lambda 37225-37255 primer p2.
KW primer; PCR; polymerase chain reaction; polysogenic; microorganism;
KW polysogen; E. coli; strain SCS-8; lambda; ci gene; SCS-8ci; pUC18;
KW N99; pUTKm; pUTKm-ci; kanamycin; transformation; s17-1/pUTKm-ci;
KW pUC18SfiI; conjugation; tetracycline; lytic phase; ss.
OS Synthetic.
PN WO9315769-A.
PD 19-AUG-1993.
PF 12-FEB-1993: U01293.
PR 14-FEB-1992; US-837031.
PR 04-SEP-1992; US-941190.
PA (STRA-) STRATAGEME.
PI Kretz PL, Short JM.
DR WPI: 93-272571/34.
PT Mutagenesis testing using transgenic non-human animals carrying
PT test DNA sequences - providing selectable reporter molecules for
PT high sensitivity and of detection
PS Disclosure; Page 90; 145pp; English.
CC The sequences given in 048145-46 are primers which were used in the
CC preparation of a polysogenic microorganism. The polysogen was
CC constructed in E. coli strain SCS-8 and a stably integrated copy of
CC the lambda ci gene was inserted into the genome to form the E. coli
CC designated SCS-8ci. The plasmid pUC18SfiI was derived from pUC18 by
CC adding two SfiI restriction sites flanking the polylinker. These two
CC primers correspond to the termini of the wild type ci gene and were
CC used on wild type lambda in the form of an E. coli lysogen extract of
CC strain N99. The amplification product was ligated into the polylinker
CC region of the plasmid pUC18SfiI and then removed from the plasmid by
CC SfiI digestion, to form a PCR product with SfiI cohesive ends. pUTKm
CC was digested with SfiI and the PCR fragment was ligated into it to
CC form pUTKm-ci. pUTKm-ci was transformed into E. coli and cultured in
CC kanamycin to select for transformants containing the plasmid, and were
CC designated s17-1/pUTKm-ci. A standard mating protocol was conducted
CC where cultures of SCS-8 and s17-1/pUTKm-ci were cocultured. Plasmid
CC pUTKm-ci was transferred by conjugation to SCS-8 cells and the
CC resulting conjugates were plated onto plates containing kanamycin and
CC tetracycline. Viable cells were designated SCS-8ci cells and are
CC polysogenic microorganisms as the expression of the ci gene prevents
CC the cells entering the lytic phase upon infection by a lysis-competent
CC lambda bacteriophage.
SQ Sequence 35 BP; 9 A; 11 C; 8 G; 7 T;

Query Match 62.4%; Score 10.6; DB 1; Length 35;
Best Local Similarity 58.8%; Pred. No. 9.2e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 gaygncarcncntaygt 17
   ||:| ||:| ||:|
Db 10 GAGCATCAGCCAAACGT 26

RESULT 15
V39428
ID V39428 standard; DNA; 35 BP.
AC V39428;
DT 21-SEP-1998 (first entry)
DE Humanised anti-HM1.24 antibody PCR primer SEQ ID NO:89.
KW Mouse; human; humanised; anti-HM1.24 antibody; myeloma; FR; CDR;
KW framework region; complementarity determining region; antigenicity;
KW PCR primer; ss.
OS Synthetic.
OS Mus sp.
OS Homo sapiens.
PN WO9814580-A1.
PD 09-APR-1998.
PF 03-OCT-1997; J03553.
PR 04-OCT-1996; JP-264756.
PA (CHUS ) CHUGAI SEIYAKU KK.
PI Kojishimura Y, Kosaka M, Ohtomo T, Ono K, Tsuchiya M,
PI Yoshimura Y;

```

```

DR WPI: 98-286421/25.
PT Humanised anti-HM1.24 antibody - for treatment of myeloma
PS Example 9; Page 146; 210pp; Japanese.
CC A humanised anti-HM1.24 antibody has been developed which comprises
CC human L and H chain C regions, and L and/or H chain V regions
CC containing material originating in mouse anti-HM1.24 antibody. The V
CC regions contain framework (FR) regions of human origin and
CC complementarity determining regions (CDR) of mouse origin, leading to
CC a reshaped humanised antibody. The C regions are human Ck (L-chain) and
CC human C gamma (especially C gamma 1) (H-chain). The FR regions of the
CC L chain V region are derived from human subtype HSG1 (e.g. from human
CC antibody RE1) and the FR regions of the H chain V region are derived
CC from human subtype HSG1 (e.g. FR1-3 from human antibody HG3 and FR4
CC from human antibody JH6). The present sequence represents a PCR primer
CC used in an example from the present invention. The antibodies are used
CC for the treatment of myeloma, especially by injection. Intravenously,
CC intramuscularly or subcutaneously. The antibodies are used at 0.01-1000
CC (especially 5-100) mg/kg body weight. The humanised antibody has low
CC antigenicity and is therefore therapeutically in humans.
SQ Sequence 35 BP; 6 A; 8 C; 11 G; 10 T;

```

```

Query Match 62.4%; Score 10.6; DB 1; Length 35;
Best Local Similarity 58.8%; Pred. No. 9.2e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 gaygncarcncntaygt 17
   ||:| ||:| ||:|
Db 9 GATGCTCAGCTGACGCT 25

```

Search completed: June 4, 2000, 16:23:57
Job time: 28893 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 08:04:16 ; Search time 4521.53 Seconds
(without alignments)
15.239 Million cell updates/sec

Title: US-09-164-714-2

Perfect score: 17
Sequence: 1 gaycncarcntaygt 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gsa1:*
83: gb_gsa2:*
84: gb_gsa3:*
85: gb_gsa4:*
86: em_gsa1:*
87: em_gsa2:*
88: em_gsa3:*
89: em_gsa4:*
90: gb_gsa5:*
91: gb_gsa6:*
92: gb_gsa7:*
93: gb_gsa8:*
94: gb_gsa9:*
95: em_gsa5:*
96: em_gsa6:*
97: em_gsa7:*
98: em_gsa8:*
99: em_gsa9:*
100: em_gsa10:*
101: em_gsa11:*
102: gb_gsa10:*
103: gb_gsa11:*
104: em_gsa12:*
105: gb_gsa12:*
106: gb_gsa13:*
107: gb_gsa14:*
108: gb_gsa15:*
109: gb_gsa16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	12.6	74.1	75	36	AA607795	AA607795 vo49f02.r
2	12.6	74.1	98	35	D63753	D63753 D63753 Huma
3	12.2	71.8	79	43	A1191144	A1191144 gel8e12.x
4	11.2	65.9	36	21	T61825	T61825 yb92a04.s1
5	11.2	65.9	89	42	A1159672	A1159672 u103a03.x
6	11.2	65.9	96	46	A1439657	A1439657 lc91g03.x
7	11	64.7	67	41	A1033279	A1033279 ow74b09.s
8	11	64.7	89	64	AM059703	AM059703 AHUTH.bs
9	10.6	62.4	46	41	A1035942	A1035942 ub49f12.r
10	10.6	62.4	49	37	AA717442	AA717442 vu19d06.r
11	10.6	62.4	67	39	AA906835	AA906835 ok73f07.s
12	10.6	62.4	67	42	A1094466	A1094466 ou87f09.s
13	10.6	62.4	85	61	A1827573	A1827573 w10e08.x
14	10.6	62.4	100	39	AA862553	AA862553 oh44b08.s
15	10.6	62.4	100	43	A1188636	A1188636 gdl5e11.x
16	10.6	62.4	100	64	AM065340	AM065340 687009h07
17	10.6	62.4	100	83	HUMAN	L48753 Homo sapien
18	10.2	60.0	49	37	AA677795	AA677795 z124d03.s
19	10.2	60.0	49	38	AA781646	AA781646 a150a01.s
20	10.2	60.0	58	40	AA958617	AA958617 ua15f12.r
21	10.2	60.0	61	23	HA6784	HA6784 y014b07.r1
22	10.2	60.0	67	44	A1273788	A1273788 qu08c02.x
23	10.2	60.0	70	29	AA177604	AA177604 mt25g10.r
24	10.2	60.0	71	61	A1869405	A1869405 tw40c01.x
25	10.2	60.0	72	37	AA675648	AA675648 vrf6f11.s
26	10.2	60.0	74	38	AA785979	AA785979 l8b10a1.f
27	10.2	60.0	76	40	AA916262	AA916262 ol63a07.s
28	10.2	60.0	82	36	AA656248	AA656248 vv12d08.r
29	10.2	60.0	82	40	AA988047	AA988047 os12g04.s
30	10.2	60.0	82	61	A1874724	A1874724 ul22e03.x
31	10.2	60.0	85	35	AA536922	AA536922 vj73h10.r
32	10.2	60.0	87	45	AJ237210	AJ237210 AJ237210
33	10.2	60.0	91	27	AA032459	AA032459 m13a11.r
34	10.2	60.0	91	30	AA219896	AA219896 mx96d11.r
35	10.2	60.0	91	74	AA214748	AA214748 uo99d06.y
36	10.2	60.0	95	37	AA701524	AA701524 z167a09.s
37	10.2	60.0	97	24	H81252	H81252 yu73g01.r1
38	10.2	60.0	97	26	W36629	W36629 mb67c04.r1
39	10.2	60.0	98	29	AA161637	AA161637 MBARG0A1
40	10	58.8	66	22	R73950	R73950 yj98h04.r1
41	10	58.8	67	62	A1877597	A1877597 fc49c10.y
42	10	58.8	84	23	H55705	H55705 CHR220644.C
43	10	58.8	85	39	AA822902	AA822902 vp18h03.r
44	10	58.8	88	35	AA574321	AA574321 nf45f03.s
45	10	58.8	91	34	AA523075	AA523075 n166c03.s

ALIGNMENTS

RESULT 1

AA607795 LOCUS vo49f02.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus EST 30-SEP-1997
 DEFINITION clone IMAGE:1053243 5', mRNA sequence. CDNA

AA607795 VERSION AA607795.1 GI:2455230

KEYWORDS

SOURCE

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

ORGANISM

REFERENCE

AUTHORS

Marras,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT On Apr 14, 1993 this sequence version replaced gi:692809.

FEATURES

source

Contact: Maria M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:584819
 putative full length read
 vector to vector length is 168
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 67.
 Location/Qualifiers
 1..75
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:1053243"
 /clone_1lb="Barstead mouse irradiated colon MPLRB7"
 /dev_stage="8 weeks"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
 from 8 week old mouse. Colon was harvested 72 hours after
 irradiation with 1400 Gys. 1st strand cDNA was primed
 with a Not I - oligo(dT) primer
 [5'-GTTTACGATCGACGTCGACGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
 T3'1]; double-stranded cDNA was ligated to Eco RI
 adaptors [ATTGGAATCTTG], digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT73
 vector. Library constructed by Bob Barstead."

BASE COUNT

ORIGIN

21 a 19 c 20 g 15 t

Query Match

Best Local Similarity 74.1%; Score 12.6; DB 36; Length 75;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 gaygcacarcctatgt 17
 |||||
 Db 18 GAGGCCCAACCCATGCT 34

RESULT 2

D63753/c

D63753 LOCUS D63753 Human chromosome 11q23 mRNA (M.Katoh) Homo sapiens CDNA
 DEFINITION clone PEXA-3, mRNA sequence.
 D63753
 ACCESSION D63753.1 GI:2326278

KEYWORDS

SOURCE

human.

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

ORGANISM

REFERENCE

AUTHORS

Katoh,M., Nakagawa,Y., Yawata,T., Kumano,S., Kobayashi,E.,

Kurimasa,A., Kugoh,H. and Oshimura,M.

Cosmids and transcribed sequences from chromosome 11q23

Jpn. J. Hum. Genet. 40 (4), 307-317 (1995)

97004453

MEDLINE

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1394933.

Contact: Motonobu Katoh

Faculty of Medicine, Department of Molecular and Cell Genetics

Tottori University

86 Nishimachi, Yonago, Tottori 683, Japan

Email: mkatoh@grape.med.tottori-u.ac.jp.

Location/Qualifiers

1..98

/organism="Homo sapiens"

FEATURES

source

1..98
 /organism="Homo sapiens"


```

/db.xref="taxon:9606"
/map="11q23"
/clone="PEX4A-3"
/clone_11p="Human chromosome 11q23 mRNA (M. Katoh)"
BASE COUNT      35 a      21 c      21 g      21 t
ORIGIN

```

Qy	1 gaggcaccatctgt	17
Dg	38 GAGCGCTACCATATGT	22

[illegible]

REFERENCE	1 (bases 1 to 79)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL	Tumor Gene Index
COMMENT	Unpublished (1997)
	On Jan 17, 1998 this sequence version replaced gi:1900121.

JOURNAL
COMMENT
Unpublished (1997)
On Jan 17, 1998 this sequence version replaced gi:1900121.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nlh.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert length: 836 Std Error: 0.00
Seq primer: 40NP from Gibco
High quality sequence stop: 1.

BASE COUNT	ORIGIN	FEATURES	SOURCE
12 a	23 c	22 t	<p>Location/Qualifiers</p> <p>1..79</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="IMAGE:173374"</p> <p>/clone_lib="Soares.fetal_lung_NBHL19W"</p> <p>/dev_stage="19 weeks"</p> <p>/lab_host="DH10B (ampicillin resistant)"</p> <p>/note="Organ: Lung; Vector: pRT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15'-TGTACCACATCTGAAGTGGAGCGCCGACATTTTTTTTTTTTTT-3' 1 double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pRT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Patima Bonaldo. This library was screened from the same fetus as the fetal heart library, Soares fetal heart NBHL19W."</p>

Query Match	71.8%;	Score 12.2;	DB 43;	Length 79;
Best Local Similarity	64.7%;	Pred. No. 1.2e+03;		
Matches 11; Conservative	3;	Mismatches 3;	Indels 0;	Gaps 0;

```
QY      1 gaygcncarcntaygt 17
          ||:| | | : | | : | |
Db      41 GATGCCCAGGCTATGT 57
```

LOCUS	36 bp	EST	14-FEB-1995
DEFINITION	yb92d04.s1 Striatogene liver (#937224) Homo sapiens cDNA clone IMAGE:7606 3' similar to gb:U1566 60S RIBOSOMAL PROTEIN L18 (HUMAN);, mRNA sequence.		
ACCESSION	T61825		
VERSION	T61825.1	GI:665068	
KEYWORDS	EST.		
SOURCE	human.		

REFERENCE	AUTHORS
1 (pages 1 to 36)	
	Hallier, L., Lennon, G., Becker, M., Bonald, M. F., Chapell, B., Chissac, S., Dietrich, N., Dubugue, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kueba, T., Laey, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M. B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, M. and Marra, M.
TITLE	Generation and analysis of 280,000 human expressed sequence tags
JOURNAL	Genome Res. 6 (9), 807-828 (1996)
MEDLINE	97044478
COMMENT	Contact: Wilson RK

High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, L1M1. This clone is available royalty-free through L1M1: contact the IMAGE Consortium (info@image.l1m1.gov) for further information. Trace considered overall poor quality
Insert Length: 693 Std Error: 0.00
Seq primer: -21m13
High quality sequence stop: 1.

FEATURES	SOURCE
Location/Qualifiers	1..36
/organism="Homo sapiens"	
/db_xref="GDB:498351"	
/db_xref="taxon:9606"	
/clone IMAGE:78606"	
/clone lib="Stratagene liver (#937224)"	
/sex="male"	
/dev_stage="49 years old"	
/lab_host="SOLR cells (kanamycin resistant)"	
/note="Organ: liver; Vector: pBluescript SK; Site: 1; EcoRI; Site: 2; XhoI; Cloned unidirectionally. Primer: Oligo dT. Hepatectomy from normal male caucasian. Average insert size: 1.1 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GATTGGCAGCG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"	
	3 a 9 c 15 g 7 t 2 others

Query Match	65.9%;	Score 11.2;	DB 21;	Length 36;
Best Local Similarity	62.5%;	Pred. No. 3.9e+03;		
Matches 10; Conservative	3;	Mismatches 3;	Indels 0;	Gaps 0

```
QY      2  aycncarcntaygt 17
          | : | | : | | : | |
Db      31  ACGCCAACCCCTACGT 16
```

RESULT

LOCUS	A1159672/c				
DEFINITION	A1159672	89 bp	mRNA	EST	02-OCT-1998
ACCESSION	U030a05.x1	Sugano mouse liver	mla	Mus musculus	cDNA clone
VERSION	IMAGE:1499504	3'	similar to gb:X04325_cds1	GAP JUNCTION	BETA-1
KEYWORDS	PROTEIN (HUMAN);	gb:M63802	Mus musculus	connexin 32 (MOUSE);	mRNA
SOURCE	A1159672				
ORGANISM	A1159672.1	GI:3692854			
REFERENCE	EST.				
AUTHORS	house mouse.				
TITLE	Mus musculus				
COMMENT	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.				
	1 (bases 1 to 89)				
	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.				
	The WashU-HMHI Mouse EST Project				
	Unpublished (1996)				
	On Jan 9, 1998 this sequence version replaced gi:936225.				

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through INLIL; contact the
IMAGE Consortium (info@image.inlil.gov) for further information.
MGI:3377108
Trace considered overall poor quality
Seq primer: custom primer used
High quality sequence stop: 1.

FEATURES

SOURCE

```

/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1499504"
/clone_1id="Sugano mouse liver ml1a"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: liver; Vector: PME18S-FL3; Site:1: DraIII
(CACCGTGG); Site:2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGAGCCCTTTTCTTTTCTTTTCTTTT] double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCCTCACTGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACCGTGG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGTCTTAAAGCTGG and 3' end
primer CGACCTGACACTCGACACA."

```

Query Match	65.9%	Score 11.2;	DB 4.2;	Length 89;
Best Local Similarity	62.5%	Pred. No.5.1e+03;		
Matches 10;	Conservative 3;	Mismatches 5;	Indels 0;	Gaps 0;
OY	1	gagcncarcrcntayg	16	
		:	:	
db	19	GATGCTCAACCTCAG	4	

RESULT 6
AI439657

LOCUS	A1439657	96 bp	mRNA	EST	28-MAR-1999
DEFINITION	cg1903.x1 NCI-CGAP, C111 Homo sapiens cDNA clone IMAGE:2073556 3				
ACCESSION	A1439657	similar to gb:X57138_rnal HISTONE H2B.2 (HUMAN);, mRNA sequence.			
VERSION	A1439657.1	GI:4305856			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 96)				
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .				
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
COMMENT	Unpublished (1997)				
	On Jun 5, 1998 this sequence version replaced gi:3189048.				

Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Ash Allizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 2241 Std Error: 0.00
Seq primer: -40UP from Gldco
High quality sequence stop: 1.

FEATURES

source

	ORGANISM="Homo sapiens"	
	/db_xref=taxon:9606"	
	/clone_image=2073556"	
	/clone_lib="NCI-CCAP_CELL"	
	/tissue_type="B-cell, chronic lymphocytic leukemia"	
	/lab_host="DH10B"	
	/note="Vector: pT73D-Pac (Pharmacia) with a modified	
	polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA	
	was primed with a Not I - oligo(dT) primer [5'	
	TGTACCACATCGTAGAGGGGAGGGGGCCTTCCTTTTTTTTTTTTTTTTTTT	
	T 3']; double-stranded cDNA was ligated to Eco RI	
	adaptors (Pharmacia), digested with Not I and cloned into	
	the Not I and Eco RI sites of the modified pT73 vector.	
	Library is normalized, and was constructed by Bento	
	Soares and M.Fatima Bonaldo."	
BASE COUNT	18 a 29 c 29 g 20 t	
ORIGIN		
Query Match	65.9%; Score 11.2; DB 46;	
Best Local Similarity	62.5%; Pred. No. 5.2e+03;	
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		
OY 1 gaycncarcntayg 16 :: :: :: :: :: :: :: :: 79 GACGCACACGCCGCTCG 94 yb		

RESULT		7				
Locus	A1033279					
Definition	A1033279	67 bp	mRNA	EST	28-AUG-1998	
Accession	U74009	S1 Soares fetal_liver.spleen_1NLS_S1 clone IMAGE:165257 3'	mRNA sequence.		Homo sapiens CDNA	
Version	A1033279					
Keywords	A1033279.1	GI:3254232				
Source	human.					
Organism	Homo sapiens Eukaryota; Metazoa;	Chordata; Craniata; Vertebrata; Mammalia;				

REFERENCE	Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS	1 (bases 1 to 67)
TITLE	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	On Jan 19, 1998 this sequence version replaced gi:2282177. Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert length: 553 Std Error: 0.00 Seq primer: -40ml3 fwd. ET from Amersham High quality sequence stop: 54. Location/Qualifiers
FEATURES	1..67
source	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:1652537" /clone_id="Soares_fetal_liver_spleen_mfLS_S1" /sex="male" /dev_stage="20 week-post conception fetus" /lab_host="DH10B (ampicillin resistant)" /note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; this is a subtracted version of the original Soares fetal liver spleen mFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACGGAAGATTTATTTAAGATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT	20 a 15 c 12 g 20 t
ORIGIN	
Query Match	64.7%; Score 11; DB 41; Length 67;
Best Local Similarity	64.7%; Pred. No. 6.2e+03;
Matches 11: Conservative	2; Mismatches 4; Indels 0; Gaps 0;
OY	1 gaygcncarcntatagt 17 : : : :
Db	31 GATGCTCGACCTTAGT 47
RESULT	8
AM059703	
LOCUS	AM059703 89 bp mRNA EST 16-FEB-2000
DEFINITION	AH07H.bsst.dnc15.ba.A020h09 DMC15 Homo sapiens cDNA, mRNA sequence.
ACCESSION	AM059703
VERSION	AM059703.1 GI:6652025
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS	1 (bases 1 to 89) Brenner,S., Williams,S.R., Vermass,E.H., Storch,T., Moon,K., McCollum,C., Mao,J.I., Kirchner,J.J., Eletz,S., Dubridge,R.B., Burcham,T. and Albrecht,G. In vitro cloning of complex mixtures of DNA on microbeads: Physical separation of differentially expressed cDNAs Proc. Natl. Acad. Sci. U.S.A. 97 (4), 1665-1670 (2000) On Jun 5, 1998 this sequence version replaced gi:3188342. Contact: Burcham TS LYNX Therapeutics, Inc. 25861 Industrial Blvd., Hayward, CA 94545, USA Tel: 510 670 9338 Fax: 510 670 9302 Email: climb@lynxgen.com
JOURNAL	
COMMENT	

```

Sequence obtained from LYNX Therapeutics. Megasort technology.
Collected from the down-regulated gate.
High quality sequence stop: 89.
Location/Qualifiers
1..89
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="DNC15"
/cell_type="monocytic leukemia"
/notes="Vector: PCR2.1; Cloning of PCR products from
micro-beads carrying 3' end of down-regulated cDNA. THP-1
cells non-induced (created with DMSO only)."
```

BASE COUNT	21 a	21 c	24 g	23 t
ORIGIN				
Query Match	64.7%	Score 11;	DB 64;	Length 89;
Best Local Similarity	64.7%	Pred. No. 6.8e+03;		
Matches 11; Conservative	2;	Mismatches 4;	Indels 0;	Gaps 0;
Db	41	GACGCACATACATACGT	57	
Oy	1	gagcncarcclataygt	17	
	:			:
RESULT 9				
AI035942/c				
LOCUS				
DEFINITION	AI035942	46 bp	mRNA	EST
	ub99f12.r1 Soares_mammary_gland_NbMNG			26-JUN-1998
	IMAGE:1381103	5'	similar to	TR:Q9969 Q9969 TAZAROTENE-INDUCED
	GENE 2.;			mRNA sequence.
ACCESSION	AI035942			
VERSION	AI035942.1	GI:3259633		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 46)			
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,			
	Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,			
	Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,			
	Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and			
	Waterston,R.			
	The WashU-HMI Mouse EST Project			
	Unpublished (1996)			
	On Sep 12, 1996 this sequence version replaced gi:1400925.			
TITLE	Contact: Marra M/Mouse EST Project			
JOURNAL	WashU-HMI Mouse EST Project			
COMMENT	Washington University School of Medicinep			
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108			
	Tel: 314 286 1800			
	Fax: 314 286 1810			
	Email: mouseest@wustl.edu			
	This clone is available royalty-free through LNL; contact the			
	IMAGE Consortium (info@image.llnl.gov) for further information.			
	MGT:903571			
	Trace considered overall poor quality			
	Seq primer: -28ml3 rev2 ET from Amersham			
	High quality sequence stop: 1.			
FEATURES	Location/Qualifiers			
SOURCE	1..46			
	/organism="Mus musculus"			
	/strain="C57BL/6J"			
	/db_xref="taxon:10090"			
	/clone="IMAGE:1381103"			
	/clone_lib="Soares_mammary_gland_NbMNG"			
	/sex="male"			
	/tissue_type="mammary gland"			
	/dev_stage="4 weeks"			
	/lab_host="DH10B"			

RESULT	13
AI827573	
LOCUS	AI827573 85 bp mRNA EST 18-DEC-1999
DEFINITION	wil0e08.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2350214 3' similar to TR:000605 O00605 PUTATIVE RAB5-INTERACTING PROTEIN [CLONE IL-57] ; , mRNA sequence.
ACCESSION	AI827573
VERSION	AI827573.1 GI:5448244
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 85)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	On Jun 5, 1998 this sequence version replaced gi:3188016. Contact: Robert Strausberg, Ph.D. Tel.: (301) 496-1550 Email: Robert.Strausberg@nih.gov This clone is available royalty-free through LNL : contact the IMAGE Consortium (infoimage.lnl.gov) for further information. Trace considered overall poor quality Insert Length: 1130 Std Error: 0.00 Seq primer: -40UP from Glibco High quality sequence stop: 1. location/Qualifiers 1..85 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2350214" /clone_1lb="Soares_NFL_T_GBC_S1" /lab_host="DH10B" /note="Organ: pooled; Vector: pTV73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI-CGAP-GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 29/480-30208/, 682632-687239/, 726408-728711/, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT	16 a 22 c 20 g 27 t
ORIGIN	
Query Match	62.4%; Score 10.6; DB 61; Length 85;
Best Local Similarity	66.7%; Pred. No. 1.2e+04;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	
OY	3 ygcncarcntaygt 17 :: :: Db 36 TGCTCATCCATATGT 50
RESULT 14	
AA862553/c	
LOCUS	AA862553 100 bp mRNA EST 24-AUG-1998
DEFINITION	oh4db08.s1 NCI-CGAP_GC4 Homo sapiens cDNA clone IMAGE:1469463 3'
ACCESSION	AA862553
VERSION	AA862553.1 GI:2955032
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - 'nucleic search, using sw model

Run on: June 4, 2000, 08:22:01 ; Search time 244.64 Seconds
(Without alignments)
9.033 Million cell updates/sec

Title: US-09-164-714-2

Perfect score: 17
Sequence: 1 gaycncarcntaygt 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/5C.COMB.seq:*
4: /cgn2_6/prodata/2/ina/5D.COMB.seq:*
5: /cgn2_6/prodata/2/ina/6-.COMB.seq:*
6: /cgn2_6/prodata/2/ina/PCrUS.COMB.seq:*
7: /cgn2_6/prodata/2/ina/backfile1.s1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11.8	69.4	68	5	US-08-469-318-99
2	11.8	69.4	68	5	US-08-469-318-100
3	11.8	69.4	68	6	PCT-US95-01185-99
4	11.8	69.4	68	6	PCT-US95-01185-100
5	11.8	69.4	68	6	PCT-US95-01185-100
6	11.8	69.4	68	6	PCT-US95-01185-100
7	11.8	69.4	68	6	PCT-US95-01185-100
8	11.8	69.4	68	6	PCT-US95-01185-100
9	11.8	69.4	68	6	PCT-US95-01185-100
10	11.8	69.4	68	6	PCT-US95-01185-100
11	11.8	69.4	68	6	PCT-US95-01185-100
12	11.8	69.4	68	6	PCT-US95-01185-100
13	11.8	69.4	68	6	PCT-US95-01185-100
14	11.8	69.4	68	6	PCT-US95-01185-100
15	11.8	69.4	68	6	PCT-US95-01185-100
16	11.8	69.4	68	6	PCT-US95-01185-100
17	11.8	69.4	68	6	PCT-US95-01185-100
18	11.8	69.4	68	6	PCT-US95-01185-100
19	11.8	69.4	68	6	PCT-US95-01185-100
20	11.8	69.4	68	6	PCT-US95-01185-100
21	11.8	69.4	68	6	PCT-US95-01185-100
22	11.8	69.4	68	6	PCT-US95-01185-100
23	11.8	69.4	68	6	PCT-US95-01185-100
24	11.8	69.4	68	6	PCT-US95-01185-100
25	11.8	69.4	68	6	PCT-US95-01185-100
26	11.8	69.4	68	6	PCT-US95-01185-100
27	11.8	69.4	68	6	PCT-US95-01185-100

28	10.2	60.0	18	2	US-08-363-240A-1108	Sequence 1108, App
29	10.2	60.0	21	1	US-08-047-413-3	Sequence 3, App1
30	10.2	60.0	21	1	US-08-047-413-4	Sequence 4, App1
31	10.2	60.0	30	1	US-08-217-210B-4	Sequence 4, App1
32	10.2	60.0	36	2	US-08-425-299A-5	Sequence 5, App1
33	10.2	60.0	50	4	US-08-556-424-5	Sequence 5, App1
34	10.2	60.0	86	4	US-08-477-527A-129	Sequence 129, App
35	10.2	60.0	86	4	US-08-477-527A-129	Sequence 129, App
36	10.2	60.0	86	4	US-08-477-527A-143	Sequence 143, App
37	10.2	60.0	86	4	US-08-477-527A-152	Sequence 152, App
38	10.2	60.0	86	5	US-08-481-710-129	Sequence 129, App
39	10.2	60.0	86	5	US-08-481-710-130	Sequence 130, App
40	10.2	60.0	86	5	US-08-481-710-143	Sequence 143, App
41	10.2	60.0	86	5	US-08-481-710-152	Sequence 152, App
42	10.2	60.0	86	6	PCT-US96-09537-129	Sequence 129, App
43	10.2	60.0	86	6	PCT-US96-09537-130	Sequence 130, App
44	10.2	60.0	86	6	PCT-US96-09537-143	Sequence 143, App
45	10.2	60.0	86	6	PCT-US96-09537-152	Sequence 152, App

ALIGNMENTS

RESULT 1
US-08-469-318-99
Sequence 99, Application US/08469318
Patent No. 6022535
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 196
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,318
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,872
FILING DATE:
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic DNA"
US-08-469-318-99

Query Match 69.4% Score 11.8 DB 5 Length 68:
Best local Similarity 66.7% Pred. No. 1.5e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 ygcncarcntaygt 17
Db 53 TGGCAGCCCTACGT 67

RESULT 2
US-08-469-318-100/c
Sequence 100, Application US/08469318
Patent No. 6022535
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 196

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,318
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,872
FILING DATE:
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic DNA"
US-08-469-318-100
```

```
Query Match          69.4%; Score 11.8; DB 5; Length 68;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      3 ygcncarcntaygt 17
      :|||:|||:|||
Db      20 TGGCGAGCCCTACT 6
```

```
RESULT      3
PCT-US95-01185-99
; Sequence 99, Application PC/TUS9501185
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01185
; FILING DATE: 02-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/192325
; FILING DATE: 14-FEB-1994
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
PCT-US95-01185-99
```

```
Query Match          69.4%; Score 11.8; DB 6; Length 68;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      3 ygcncarcntaygt 17
      :|||:|||:|||
Db      53 TGGCGAGCCCTACT 67
```

```
RESULT      4
```

```
PCT-US95-01185-100/C
; Sequence 100, Application PC/TUS9501185
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01185
; FILING DATE: 02-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/192325
; FILING DATE: 14-FEB-1994
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
PCT-US95-01185-100
```

```
Query Match          69.4%; Score 11.8; DB 6; Length 68;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      3 ygcncarcntaygt 17
      :|||:|||:|||
Db      20 TGGCGAGCCCTACT 6
```

```
RESULT      5
US-08-852-807-28/C
; Sequence 28, Application US/08852807
; Patent No. 5861298
; GENERAL INFORMATION:
; APPLICANT: Debouck, Christine
; APPLICANT: Drake, Fred
; APPLICANT: Gowen, Maxine
; APPLICANT: Rood, Julie
; APPLICANT: Hastings, Gregg
; APPLICANT: Adams, Mark
; APPLICANT: Fraser, Claire
; APPLICANT: Lee, No. 5861298man
; APPLICANT: Kirkness, Ewen
; APPLICANT: Blake, Judith
; APPLICANT: Fitzgerald, Lisa
; TITLE OF INVENTION: CATHEPSIN K GENE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,807
; FILING DATE: 07-MAY-1997
; CLASSIFICATION: 435
```

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER: 60/019,942
3      FILING DATE: 14-JUNE-1996
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER: 60/020,273
6      FILING DATE: 17-JUNE-1996
7      PRIOR APPLICATION DATA:
8      APPLICATION NUMBER: 60/026,273
9      FILING DATE: 26-AUG-1996
10     ATTORNEY/AGENT INFORMATION:
11     NAME: Han, William T
12     REGISTRATION NUMBER: 34,344
13     REFERENCE/DOCKET NUMBER: ATG50006-2
14     TELECOMMUNICATION INFORMATION:
15     TELEPHONE: 610-270-5219
16     TELEFAX: 610-270-5090
17     TELEX:
18     INFORMATION FOR SEQ ID NO: 28:
19     SEQUENCE CHARACTERISTICS:
20     LENGTH: 21 base pairs
21     TYPE: nucleic acid
22     STRANDEDNESS: single
23     TOPOLOGY: linear
24     MOLECULE TYPE: Genomic DNA
25     HYPOTHEICAL: NO
26     ANTI-SENSE: NO
27     FRAGMENT TYPE:
28     ORIGINAL SOURCE:
29     JS-08-852-807-28

```

Qy	1	gagcgcncarcntatg	17	64.7%;	Score 11;	DB 3;	Length 21;
		: :					
db	18	gATGcCTACCCATATGT	2	Best Local Similarity	64.7%;	Pred. No. 4.2e+02;	
				Matches 11;	Conservative 2;	Mismatches 4;	Indels 0;
						Gaps 0;	

US-08-852-807-29
Sequence 29, Application US/08852807
Patent No. 5861238
GENERAL INFORMATION:
APPLICANT: Debouck, Christine
APPLICANT: Drake, Fred
APPLICANT: Gowen, Maxine
APPLICANT: Rood, Julie
APPLICANT: Haslings, Gregg
APPLICANT: Adams, Mark
APPLICANT: Fraser, Claire
APPLICANT: Lee, No. 5861298man
APPLICANT: Kirkness, Ewen
APPLICANT: Blake, Judith
APPLICANT: Fitzgerald, Lisa
TITLE OF INVENTION: CATHEPSIN K GENE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEO Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,807
FILING DATE: 07-MAY-1997

```

1 CLASSIFICATION: 435
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: 60/019, 942
4 FILING DATE: 14-JUNE-1996
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: 60/020, 273
7 FILING DATE: 17-JUNE-1996
8 PRIOR APPLICATION DATA:
9 APPLICATION NUMBER: 60/026, 273
10 FILING DATE: 26-AUG-1996
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Han, William T
13 REGISTRATION NUMBER: 34, 344
14 REFERENCE/DOCKET NUMBER: ATG50006-2
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 610-270-5219
17 TELEFAX: 610-270-5090
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
10
```

Query Match	64.7%;	Score 11;	DB 3;	Length 21;
Best Local Similarity	64.7%;	Pred. No. 4.2e+02;		
Matches 11; Conservative	2;	Mismatches 4;	Indels 0;	Gaps 0;

QY	1	gaygcncarcntaygt	17
		:	
Db	4	GATGCTACCATATGT	20

RESULT 7
 US-08-850-049-18
 Sequence 18 Application US/088550049
 Patent No. 5965726
 GENERAL INFORMATION:
 APPLICANT:
 APPLICANT:
 APPLICANT:
 TITLE OF INVENTION: METHOD OF ELIMINATING
 TITLE OF INVENTION: INHIBITORY/INSTABILITY REGIONS OF RNA
 NUMBER OF SEQUENCES: 130
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/850,049
 FILING DATE: 02-MAY-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/050,478
 FILING DATE: 26-OCT-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/02908
FILING DATE: 29-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/858,747
FILING DATE: 27-MAR-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MORRIS, MARY J.
REGISTRATION NUMBER: 34,398
REFERENCE/DOCKET NUMBER: 2026-4006US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)751-6849
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
US-08-850-049-18

Query Match 64.7%; Score 11; DB 4; Length 50;
Best Local Similarity 64.7%; Pred. No. 4.4e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 gaygcncarcctaygt 17
||| ||: || ||: ||
Db 19 GAAGCAGCAGCCACTACT 35

RESULT 8
US-08-050-478-18
Sequence 18, Application US/08050478
Patent No. 5972596
GENERAL INFORMATION:
APPLICANT:
APPLICANT:
TITLE OF INVENTION: METHOD OF ELIMINATING
TITLE OF INVENTION: INHIBITORY/INSTABILITY REGIONS OF mRNA
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/050,478
FILING DATE: 26-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/02908
FILING DATE: 29-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/858,747
FILING DATE: 27-MAR-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MORRIS, MARY J.
REGISTRATION NUMBER: 34,398
REFERENCE/DOCKET NUMBER: 2026-4006US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)751-6849

TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
US-08-050-478-18

Query Match 64.7%; Score 11; DB 4; Length 50;
Best Local Similarity 64.7%; Pred. No. 4.4e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 gaygcncarcctaygt 17
||| ||: || ||: ||
Db 19 GAAGCAGCAGCCACTACT 35

RESULT 9
PCT-US94-06079-41/c
Sequence 41, Application PC/TUS9406079
GENERAL INFORMATION:
APPLICANT: American Cyanamid Company,
TITLE OF INVENTION: Gene Insertion by Direct Ligation In
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470-8426
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06079
FILING DATE: 27-MAY-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31969-00\PCF
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US94-06079-41

Query Match 64.7%; Score 11; DB 6; Length 63;
Best Local Similarity 64.7%; Pred. No. 4.5e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 gaygcncarcctaygt 17
||| ||: || ||: ||
Db 24 GAGCCAGCAGCCAGCT 8

RESULT 10
US-08-009-265-21/c

COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,710
FILING DATE: 7-JUNE-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/931,473
FILING DATE: 17-AUGUST-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,991
FILING DATE: 8-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
ATTORNEY/AGENT INFORMATION:
NAME: Diane H. McLearn
REGISTRATION NUMBER: 33,960
REFERENCE/DOCKET NUMBER: NEX41-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 86 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All pyrimidines are 2'-F
US-08-481-710-94

Query Match 64.7%; Score 11; DB 5; Length 86;
Best Local Similarity 64.7%; Pred. No. 4.6e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 gaycncarcntaygt 17
||| ||| : ||| |||
DB 65 GAAGCTCAATCATATCT 49
RESULT 13
PCT-US96-09537-94/C
Sequence 94, Application PC/TUS9609537
GENERAL INFORMATION:
APPLICANT: NEXSTAR PHARMACEUTICALS, INC.
APPLICANT: DIANE TASSETT
APPLICANT: NIKOS PAGRATIS
APPLICANT: SUMEDHA JAYASENA
APPLICANT: LARRY GOLD
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS OF
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Swanson and Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Denver
STATE: Colorado
COUNTRY: USA

ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09537
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,829
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/481,710
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX41/PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 86 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All pyrimidines are 2'-F
PCT-US96-09537-94

Query Match 64.7%; Score 11; DB 6; Length 86;
Best Local Similarity 64.7%; Pred. No. 4.6e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 gaycncarcntaygt 17
||| ||| : ||| |||
DB 65 GAAGCTCAATCATATGT 49
RESULT 14
US-08-009-265-39/C
Sequence 39, Application US/08009265
Patent No. 5547871
GENERAL INFORMATION:
APPLICANT: Black Dr., Bruce C.
APPLICANT: Summers Dr., Max D.
TITLE OF INVENTION: Heterologous Signal Sequences For
TITLE OF INVENTION: Secretion Of Insect Controlling Proteins
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESSES:
ADDRESSEE: American Cyanamid Company
STREET: 1937 West Main Street, P.O. Box 60
CITY: Stamford
STATE: CT
COUNTRY: USA
ZIP: 06904-0060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/009,265
FILING DATE: 19930125
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gordon M., Alan M.

? REGISTRATION NUMBER: 30637
 ? REFERENCE/DOCKET NUMBER: 31668-00
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: 203-321-2719
 ? TELEFAX: 203-321-2971
 ? TELE: 710-474-4059
 ? INFORMATION FOR SEQ ID NO: 39:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 90 base pairs
 ? TYPE: NUCLEIC ACID
 ? STRANDEDNESS: single
 ? TOPOLOGY: linear
 ? MOLECULE TYPE: DNA (genomic)
 ? HYPOTHEetical: NO
 ? ANTI-SENSE: NO
 ? ORIGINAL SOURCE:
 ? ORGANISM: Drosophila melanogaster
 ? US-08-009-265-39

[illegible]

US-08-477-527A-103
Sequence 103, Application US/08477527A
Patent No. 5972599
GENERAL INFORMATION:
APPLICANT: DIANE TASSER
APPLICANT: NIKOS PAGRATIS
APPLICANT: SUMEDHA JAYASENA
APPLICANT: LARRY GOLD
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS
TITLE OF INVENTION: OF CYTOKINES
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson and Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch, 1.44 MB storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,527A
FILING DATE: 7-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/931,473
FILING DATE: 17-AUGUST-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,591
FILING DATE: 8-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
ATTORNEY/AGENT INFORMATION:
NAME: Diane H. McCleathn

```

1  REGISTRATION NUMBER: 33,960
2  REFERENCE/DOCKET NUMBER: NEX11-1
3  TELECOMMUNICATION INFORMATION:
4  TELEPHONE: (303) 793-3333
5  TELEFAX: (303) 793-3433
6  INFORMATION FOR SEQ. ID NO: 103:
7  SEQUENCE CHARACTERISTICS:
8  LENGTH: 80 base pairs
9  TYPE: nucleic acid
10 STRANDEDNESS: single
11 TOPOLOGY: linear
12 MOLECULE TYPE: RNA
13 FEATURE:
14 OTHER INFORMATION: All pyrimidines are 2'-F-
15 OTHER INFORMATION: modified
16 US-08-477-527A-103

```

Query Match	63.5%	Score 10.8	DB 4	Length 80
Best Local Similarity	57.1%	Pred. NO. 6e+02		
Matches	8	Conservative	4	Mismatches 2
				Indels 0
				Gaps 0
Qy	2	aygcncarcntay	15	
		: : : : : :		
Db	16	ACGCUCACCGGUAU	29	

Search completed: June 4, 2000, 16:09:37
Job time: 28056 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:23 ; Search time 1236.38 Seconds

(without alignments)
-13.376 Million cell updates/sec

Title: US-09-164-714-3

Perfect score: 17

Sequence: 1 tgytngcrtcnacytg 17

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_cm:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_ov:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl1:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pt4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*

45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pt5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	12.6	74.1	86	5	AR042759	AR042759 Sequence
2	12.6	74.1	86	5	I65650	I65650 Sequence 10
3	12.6	74.1	86	5	I67882	I67882 Sequence 10
4	12.6	74.1	86	5	I90103	I90103 Sequence 10
5	12.2	71.8	48	5	E07909	E07909 PCR primer
6	12.2	71.8	98	5	A59486	A59486 Sequence 36
7	11.2	65.9	50	5	I42237	I42237 Sequence 50
8	11.2	65.9	53	5	AR051504	AR051504 Sequence 7
9	11.2	65.9	54	5	A43617	A43617 Sequence 7
10	11.2	65.9	54	5	AR040961	AR040961 Sequence
11	11.2	65.9	57	5	AR040962	AR040962 Sequence
12	11.2	65.9	57	5	AR051866	AR051866 Sequence
13	11.2	65.9	57	5	AR051867	AR051867 Sequence
14	11.2	65.9	73	8	ATSAABE08	U19128 Arabidopsis
15	11.2	65.9	84	3	CATFES13	M16699 Feline c-fe
16	11.2	65.9	93	7	EA28SD1	Z54117 E.album 28S
17	11.2	65.9	93	8	BAU35417	U35417 Beauveria a
18	11.2	65.9	96	5	E07904	E07904 DNA fragmen
19	11.2	65.9	99	9	HUMTHA	M20911 Human tyros
20	11	64.7	21	3	AB005886	AB005886 Bos tauru
21	11	64.7	31	5	A45726	A45726 Sequence 5
22	11	64.7	80	34	DROTRO06	L00359 D.melanogas
23	10.6	62.4	20	5	A82484	A82484 Sequence 22
24	10.6	62.4	22	5	I46451	I46451 Sequence 43
25	10.6	62.4	22	5	I46452	I46452 Sequence 43
26	10.6	62.4	22	5	I46453	I46453 Sequence 43
27	10.6	62.4	24	5	A27856	A27856 Mutagenese
28	10.6	62.4	27	5	I12386	I12386 Sequence 9
29	10.6	62.4	30	5	AR004750	AR004750 Sequence
30	10.6	62.4	30	5	AR008236	AR008236 Sequence
31	10.6	62.4	30	5	I77020	I77020 Sequence 80
32	10.6	62.4	30	5	I81015	I81015 Sequence 80
33	10.6	62.4	30	5	I81111	I81111 Sequence 80
34	10.6	62.4	32	5	A83637	A83637 Sequence 4
35	10.6	62.4	36	5	I17695	I17695 Sequence 12
36	10.6	62.4	42	5	A05115	A05115 Oligonucleo
37	10.6	62.4	43	12	MMQAL1RNA	X60888 M.musculus
38	10.6	62.4	50	14	SYNP17XORC	K00453 Chimeric pl
39	10.6	62.4	51	12	U92176	U92176 Mus musculi
40	10.6	62.4	56	5	E03956	E03956 DNA encodin
41	10.6	62.4	57	12	AF224185	AF224185 Mus muscu
42	10.6	62.4	59	5	E00072	E00072 Partial cdn
43	10.6	62.4	75	5	I12390	I12390 Sequence 13
44	10.6	62.4	81	16	AF018384	AF018384 Hepatitis
45	10.6	62.4	81	16	AF018385	AF018385 Hepatitis

ALIGNMENTS

RESULT	1	AR042759	86 bp	DNA	PAT	29-SEP-1999
LOCUS		AR042759	Sequence 168 from patent US 5811533.			
DEFINITION		AR042759				
ACCESSION		AR042759				
VERSION		AR042759.1	GI:5963255			
KEYWORDS						
SOURCE		Unknown.				
ORGANISM		Unknown.				
REFERENCE		Unclassified.				
AUTHORS		1 (bases 1 to 86)				
TITLE		Gold,L. and Janjic,N.				
JOURNAL		High-affinity oligonucleotide ligands to vascular endothelial				
FEATURES		growth factor (VEGF)				
source		Patent: US 5811533-A 168 22-SEP-1998;				
		Location/Qualifiers				
		1..86				
BASE COUNT		21 a	22 c	24 g	18 t	1 others
ORIGIN		/organism="unknown"				
Query Match		Best Local Similarity	74.1%;	Score 12.6;	DB 5;	Length 86;
Matches	12;	Conservative	70.6%;	Pred. No. 3.3e+03;	Mismatches	3;
					Indels	0;
					Gaps	0;
OY	1	tygttngertcnaactg	17			
Db	50	TGTTTCGGCTCCACTTG	66			
RESULT	2	165650	86 bp	DNA	PAT	07-OCT-1997
LOCUS		165650	Sequence 10 from patent US 5668264.			
DEFINITION		165650				
ACCESSION		165650.1	GI:2482220			
VERSION						
KEYWORDS		Unknown.				
SOURCE		Unknown.				
ORGANISM		Unclassified.				
REFERENCE		1 (bases 1 to 86)				
AUTHORS		Janjic,N. and Gold,L.				
TITLE		High affinity PDGF nucleic acid ligands				
JOURNAL		Patent: US 5668264-A 10 16-SEP-1997;				
FEATURES		Location/Qualifiers				
source		1..86				
		/organism="unknown"				
BASE COUNT		22 a	21 c	25 g	18 t	
ORIGIN						
Query Match		Best Local Similarity	74.1%;	Score 12.6;	DB 5;	Length 86;
Matches	12;	Conservative	70.6%;	Pred. No. 3.3e+03;	Mismatches	3;
					Indels	0;
					Gaps	0;
OY	1	tygttngertcnaactg	17			
Db	50	TGTTTCGGCTCCACTTG	66			
RESULT	3	167882	86 bp	DNA	PAT	04-FEB-1998
LOCUS		167882	Sequence 10 from patent US 5674685.			
DEFINITION		167882				
ACCESSION		167882.1	GI:2830004			
VERSION						
KEYWORDS		Unknown.				
SOURCE		Unknown.				
ORGANISM		Unclassified.				
REFERENCE		1 (bases 1 to 86)				

AUTHORS	Janlic,N. and Gold,L.
TITLE	High affinity PDGF nucleic acid ligands
JOURNAL	Patent: US 5674685-A 10 07-Oct-1997;
FEATURES	Location/Qualifiers
source	1..86
BASE COUNT	/organism="unknown"
ORIGIN	22 a 21 c 25 g 18 t
Query Match	74.1%; Score 12.6; DB 5; Length 86;
Best Local Similarity	70.6%; Pred. No.3.3e+03;
Matches	12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY	1 tgytngertcnaacytg 17 : : Db 50 TGGTTAGCGCTCCACTTG 66
RESULT	4
LOCUS	190103 86 bp DNA PAT 10-AUG-1998
DEFINITION	Sequence 10 from patent US 5723594.
ACCESSION	I90103
VERSION	I90103.1 GI:3410043
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 86)
AUTHORS	Janlic,N. and Gold,L.
TITLE	High affinity PDGF nucleic acid ligands
JOURNAL	Patent: US 5723594-A 10 03-MAR-1998;
FEATURES	Location/Qualifiers
source	1..86
BASE COUNT	/organism="unknown"
ORIGIN	22 a 21 c 25 g 18 t
Query Match	74.1%; Score 12.6; DB 5; Length 86;
Best Local Similarity	70.6%; Pred. No.3.3e+03;
Matches	12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY	1 tgytngertcnaacytg 17 : : Db 50 TGGTTAGCGCTCCACTTG 66
RESULT	5
LOCUS	E07909/c 48 bp DNA PAT 29-SEP-1997
DEFINITION	PCR primer.
ACCESSION	E07909
KEYWORDS	E07909.1 GI:2176041
SOURCE	JP 1994205691-A/18. unidentified. unclassified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 48)
AUTHORS	Hayashi,N.
TITLE	PRODUCTION OF MONOCLONAL ANTIBODY
JOURNAL	Patent: JP 1994205691-A 18 26-JUL-1994; SAMADELI TECHNOL:KK, HAYASHI NAKANOSHU
COMMENT	OS None OC Artificial sequences. PN JP 1994205691-A/18 PD 26-JUL-1994 PF 08-JAN-1993 JP 1993001831 PI HAYASHI NAKANOSHU PC C12P21/08,C12N15/62//C12M15/13,(C12P21/08,C12R1.19); CC strandedness: Single; topology: Linear;

CC hypothetical: No; Location/Qualifiers
FH Key
FT source 1..48 /organism='Artificial sequences',
source 1..48 /organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 16 a 7 c 14 g 11 t
ORIGIN

Query Match 71.8%; Score 12.2; DB 5; Length 48;
Best Local Similarity 64.7%; Pred. No. 5.8e+03;
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacytg 17
11:11 11:11 11:11
Db 37 TGTATTACGTCACTG 21

RESULT 6
LOCUS A59486 98 bp DNA PAT 06-MAR-1998
DEFINITION Sequence 36 from Patent WO9705234.
ACCESSION A59486
VERSION A59486.1 GI:3714798
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified

REFERENCE 1 (bases 1 to 98)
AUTHORS Chamberlain,S., Pook,M.A., Doudney,C., William,E., Hillermann,R.,
TITLE Garcia-Valdecasas,J.J. and C.
JOURNAL GENE FOR FRIEDREICH'S ATAXIA
PATENT: WO 9705234-A 36 13-FEB-1997;
IMPERIAL COLLEGE (GB)
FEATURES
source 1..98
Location/Qualifiers
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 30 a 23 c 22 g 23 t
ORIGIN

Query Match 71.8%; Score 12.2; DB 5; Length 98;
Best Local Similarity 64.7%; Pred. No. 5.8e+03;
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacytg 17
11:11 11:11 11:11
Db 70 TGTGTGGCGTCTCCTG 54

RESULT 7
LOCUS 142237 50 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 50 from patent US 5629153.
ACCESSION 142237
VERSION 142237.1 GI:2467732
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 50)
AUTHORS Undee,M.S.
TITLE Use of DNA-dependent RNA polymerase transcripts as reporter
molecules for signal amplification in nucleic acid hybridization
assays
JOURNAL Patent: US 5629153-A 50 13-MAY-1997;
FEATURES
source 1..50
Location/Qualifiers

BASE COUNT 8 a 11 c 13 g 18 t
ORIGIN

Query Match 65.9%; Score 11.2; DB 5; Length 50;
Best Local Similarity 62.5%; Pred. No. 2.4e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacyt 16
11:11 11:11 11:11
Db 21 TGTATGATCTCTCT 36

RESULT 8
LOCUS AR051504/c 53 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 74 from patent US 5830670.
ACCESSION AR051504
VERSION AR051504.1 GI:5974868
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 53)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's
disease
JOURNAL Patent: US 5830670-A 74 03-NOV-1998;
FEATURES
source 1..53
Location/Qualifiers
/organism='unknown'
BASE COUNT 13 a 11 c 13 g 16 t
ORIGIN

Query Match 65.9%; Score 11.2; DB 5; Length 53;
Best Local Similarity 62.5%; Pred. No. 2.4e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 gytngcrtcnacytg 17
11:11 11:11 11:11
Db 32 GCTATGATCACTTG 17

RESULT 9
LOCUS A43617 54 bp DNA PAT 06-MAR-1997
DEFINITION Sequence 7 from Patent WO9507987.
ACCESSION A43617
VERSION A43617.1 GI:2298819
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified

REFERENCE 1 (bases 1 to 54)
AUTHORS Budlolt,M., De,W.C., Colau,D. and Roux-Salembien,P.
TITLE NOVEL PROTEINS/POLYPEPTIDES AND COTRANSCRIPTION PLASMIDS AND LIVE
RECOMBINANT CARRIERS THEREFOR
JOURNAL Patent: WO 9507987-A 7 23-MAR-1995;
COMMENT SOLVAY (BE)
Other publication NO 961086 960509
Other publication ZA 940687 950627
Other publication AU 7615894 950403
Other publication GB 2282601 950412.
FEATURES
source 1..54
Location/Qualifiers
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 16 a 8 c 15 g 15 t
ORIGIN

Query Match 65.9%; Score 11.2; DB 5; Length 54;
Best Local Similarity 62.5%; Pred. No. 2.4e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 tgytngcrtcnacyt 17
11:1 11:1 11:1 11:1
Db 34 GTTGTGATCCACTG 19

RESULT 10
AR040961 54 bp DNA PAT 29-SEP-1999
LOCUS AR040961
DEFINITION Sequence 60 from patent US 5811238.
ACCESSION AR040961
VERSION AR040961.1 GI:5961457
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 54)
AUTHORS Stemmer, W.P.C. and Cramerl, A.
TITLE Methods for generating polynucleotides having desired
characteristics by iterative selection and recombination
JOURNAL Patent: US 5811238-A 60 22-SEP-1998;
FEATURES Location/Qualifiers
source 1..54
/organism="unknown"
BASE COUNT 9 a 18 c 15 g 12 t
ORIGIN

Query Match 65.9%; Score 11.2; DB 5; Length 54;
Best Local Similarity 62.5%; Pred. No. 2.4e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacyt 16
11:1 11:1 11:1 11:1
Db 7 TGCTGCGCATCCACCT 22

RESULT 11
AR040962 57 bp DNA PAT 29-SEP-1999
LOCUS AR040962
DEFINITION Sequence 61 from patent US 5811238.
ACCESSION AR040962
VERSION AR040962.1 GI:5961458
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 57)
AUTHORS Stemmer, W.P.C. and Cramerl, A.
TITLE Methods for generating polynucleotides having desired
characteristics by iterative selection and recombination
JOURNAL Patent: US 5811238-A 61 22-SEP-1998;
FEATURES Location/Qualifiers
source 1..57
/organism="unknown"
BASE COUNT 13 a 17 c 10 g 17 t
ORIGIN

Query Match 65.9%; Score 11.2; DB 5; Length 57;
Best Local Similarity 62.5%; Pred. No. 2.4e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacyt 16
11:1 11:1 11:1 11:1
Db 7 TGCTGCGCATCCACCT 22

RESULT 12
AR051866

LOCUS AR051866 57 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 58 from patent US 5830721.
ACCESSION AR051866
VERSION AR051866.1 GI:5975230
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 57)
AUTHORS Stemmer, W.P.C. and Cramerl, A.
TITLE DNA mutagenesis by random fragmentation and reassembly
JOURNAL Patent: US 5830721-A 58 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..57
/organism="unknown"
BASE COUNT 9 a 18 c 15 g 12 t 3 others
ORIGIN

Query Match 65.9%; Score 11.2; DB 5; Length 57;
Best Local Similarity 62.5%; Pred. No. 2.4e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacyt 16
11:1 11:1 11:1 11:1
Db 7 TGCTGCGCATCCACCT 22

RESULT 13
AR051867 57 bp DNA PAT 29-SEP-1999
LOCUS AR051867
DEFINITION Sequence 59 from patent US 5830721.
ACCESSION AR051867
VERSION AR051867.1 GI:5975231
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 57)
AUTHORS Stemmer, W.P.C. and Cramerl, A.
TITLE DNA mutagenesis by random fragmentation and reassembly
JOURNAL Patent: US 5830721-A 59 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..57
/organism="unknown"
BASE COUNT 13 a 17 c 10 g 17 t
ORIGIN

Query Match 65.9%; Score 11.2; DB 5; Length 57;
Best Local Similarity 62.5%; Pred. No. 2.4e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacyt 16
11:1 11:1 11:1 11:1
Db 7 TGCTGCGCATCCACCT 22

RESULT 14
AT5ABRE08 73 bp DNA PLN 21-MAR-1995
LOCUS AT5ABRE08
DEFINITION Arabidopsis thaliana SABRE gene, exon 8.
ACCESSION U19128
VERSION U19128.1 GI:719283
KEYWORDS
SEGMENT 8 of 14
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
REFERENCE 1 (bases 1 to 73)

AUTHORS Aeschbacher, R.A., Hauser, M.T., Feldmann, K.A. and Benfey, P.N.
TITLE The SABRE gene is required for normal cell expansion in Arabidopsis
JOURNAL Genes Dev. 9 (3), 330-340 (1995)
MEDLINE 951/2383
REFERENCE 2 (bases 1 to 73)
AUTHORS Benfey, P.N.
TITLE Direct Submision
JOURNAL Submitted (27-DEC-1994) Phillip N. Benfey, Biology, New York
University, 1009 Main Building, 100 Washington Square East, New
York, NY 10003, USA

FEATURES

source
1..73
/organism="Arabidopsis thaliana"
/isolate="WS ecotype"
/db_xref="taxon:3702"
/clone="SABRE"
/chromosome="1"
1..73
/gene="SABRE"
/number=8
BASE COUNT 21 a 17 c 18 g 17 t
ORIGIN

Query Match 65.9%; Score 11.2; DB 3; Length 84;
Best Local Similarity 71.4%; Pred. No. 2.4e+04;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgytngcrtcnac 14
||:| | | | | | | | | |
Db 30 TGCTTGCAATCCAC 17

RESULT 15

CATFES13 84 bp DNA MAM 27-APR-1993
LOCUS Feline c-fes/fps proto-oncogene-encoded protein, exon 13.
DEFINITION
ACCESSION M16699
VERSION M16699.1 GI:163840

KEYWORDS fes/fps cellular oncogene; proto-oncogene.
SEGMENT 13 of 19

SOURCE Feline DNA.
ORGANISM Felis catus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Carnivora; Fissipedia; Felidae; Felis.
AUTHORS Roebroek, A.J.M., Schalken, J.A., Onnekink, C., Bloemers, H.P.J. and
Van de Ven, W.J.

TITLE Structure of the feline c-fes/fps proto-oncogene: Genesis of a
retroviral oncogene

JOURNAL J. Virol. 61, 2009-2016 (1987)
MEDLINE 87198954

FEATURES
source
1..84
Location/Qualifiers

prim_transcript <1..>84
/db_xref="taxon:9685"
/organism="Felis catus"

Intron <1..>20
/note="c-fes/fps proto-oncogene mRNA and Intron"

exon 21..74
/note="c-fes/fps Intron L"

Intron 75..>84
/number=13

BASE COUNT 17 a 23 c 29 g 15 t
ORIGIN About 360 bp after segment 12.

Query Match 65.9%; Score 11.2; DB 3; Length 84;
Best Local Similarity 71.4%; Pred. No. 2.4e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 gytngcrtcnac 17

Db 38 GTTAGCACCCTTG 23
|:| | | | | | | | | |

Search completed: June 4, 2000, 16:05:26
Job time: 27878 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:23:57 ; Search time 322.35 Seconds
(without alignments)
13.195 Million cell updates/sec

Title: US-09-164-714-3

Perfect score: 17
Sequence: 1 tgytngcrtnacgtg 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12.6	74.1	86	1	T65226
2	12.6	74.1	86	1	T88603
3	12.2	71.8	31	1	O90000
4	12.2	71.8	98	1	T84384
5	11.8	69.4	34	1	O90002
6	11.2	65.9	50	1	O12936
7	11.2	65.9	50	1	T74386
8	11.2	65.9	54	1	O97417
9	11.2	65.9	54	1	V54949
10	11.2	65.9	57	1	O97418
11	11.2	65.9	57	1	V54950
12	11.2	65.9	74	1	T75059
13	11.2	65.9	31	1	O99185
14	10.8	63.5	22	1	V21071
15	10.6	62.4	20	1	X16628
16	10.6	62.4	22	1	X10014
17	10.6	62.4	24	1	O21863
18	10.6	62.4	27	1	O24878
19	10.6	62.4	29	1	O89996
20	10.6	62.4	32	1	V45582
21	10.6	62.4	36	1	O23643
22	10.6	62.4	43	1	V84742
23	10.6	62.4	45	1	O94185
24	10.6	62.4	48	1	O76356
25	10.6	62.4	50	1	O76357
26	10.6	62.4	50	1	O76350
27	10.6	62.4	62	1	V33733
28	10.6	62.4	63	1	O41561
29	10.6	62.4	65	1	N20038
30	10.6	62.4	75	1	O24882
31	10.6	62.4	80	1	O46873
32	10.6	62.4	87	1	O68700
33	10.6	62.4	90	1	N92126
34	10.6	62.4	97	1	T57696

C 35	10.6	62.4	100	1	T30882	Primer 12 for 95 k
C 36	10.2	60.0	18	1	O28471	Thrombin aptamer p
C 37	10.2	60.0	18	1	T93171	5'-terminal flanki
C 38	10.2	60.0	18	1	V18757	Primer LE407 for c
C 39	10.2	60.0	18	1	V60398	5' PCR primer used
C 40	10.2	60.0	20	1	O82518	Chromosome 11 (loc
C 41	10.2	60.0	20	1	V03256	Homo sapiens mutan
C 42	10.2	60.0	21	1	X00302	LacZ specific PCR
C 43	10.2	60.0	22	1	T99764	Bacillus thuringie
C 44	10.2	60.0	22	1	T99735	Bacillus thuringie
C 45	10.2	60.0	25	1	V03257	Homo sapiens mutan

ALIGNMENTS

RESULT 1	T65226
ID	T65226 standard; DNA; 86 BP.
AC	T65226;
DT	10-SEP-1997 (first entry)
DE	Platelet derived growth factor binding ligand 8.
KM	Platelet derived growth factor; PDGF; binding ligand;
KM	Identification: SELEX: anti-mitogenic; inhibition; cell;
KM	Systematic Evolution of Ligands by EXponential enrichment;
KM	epithelial; proliferation; diagnosis; treatment;
KM	neointimal lesion; arterial injury; ss.
OS	Synthetic.
PN	WO9638579-A1.
PD	05-DEC-1996.
PF	30-MAY-1996; U08014.
PR	02-JUN-1995; US-458424.
PR	02-JUN-1995; US-458423.
PR	05-JUN-1995; US-465594.
PR	05-JUN-1995; US-465591.
PR	07-JUN-1995; US-479783.
PR	07-JUN-1995; US-479725.
PR	20-MAR-1996; US-618693.
PA	(NEXS-) NEXSTAR PHARM INC.
PI	Gold L, Janjic N, Pagratlis N, Ringquist S, Toothman PJ;
DR	WPI; 97-034387/03.
PT	Identification of nucleic acid ligands to TGF-beta, PDGF and bKGF -
PT	using SELEX, used in the diagnosis and treatment of proliferative
PT	disorders
PS	Claim 39; Page 127; 209pp; English.
CC	The present sequence, a platelet derived growth factor (PDGF)
CC	binding ligand, was identified by Systematic Evolution of Ligands
CC	by EXponential enrichment (SELEX). Briefly a candidate mixture of
CC	nucleic acids was contacted with PDGF, and nucleic acids having an
CC	increased affinity to PDGF partitioned from the remainder of the
CC	mixture. The partitioned nucleic acids were then amplified to yield
CC	a mixture of nucleic acids enriched for sequences with higher
CC	affinity and specificity for binding to PDGF. The ligand is
CC	anti-mitogenic and may be used to inhibit epithelial cell
CC	proliferation, or in the diagnosis and treatment of PDGF mediated
CC	pathological conditions, e.g. to inhibit the development of
CC	neointimal lesions following arterial injury.
SQ	Sequence 86 BP; 22 A; 21 C; 25 G; 18 T;
Query Match	74.1%; Score 12.6; DB 1; Length 86;
Best Local Similarity	70.6%; Pred. No. 1.9e+02;
Matches 12; Conservative	2; Mismatches 3; Indels 0; Caps 0;
QY	1 tgytngcrtnacgtg 17
DB	50 TGGTACGTCACCTG 66
RESULT 2	T88603
ID	T88603 standard; DNA; 86 BP.
AC	T88603;

DT 18-JAN-1998 (first entry)
 DE VEGF nucleic acid ligand used to treat angiogenesis-mediated diseases.
 KM Nucleic acid ligand; secretory phospholipase A2; inflammation;
 KW anti-inflammatory; rheumatoid arthritis; septic shock; VEGF;
 KM vascular endothelial growth factor; angiogenesis; psoriasis;
 KM metastasis; neoplasia; haemangioma; human immunodeficiency virus;
 KM HIV diagnosis; treatment; specific; high affinity; ss.
 OS Synthetic.
 PN WO9627604-A1.
 PD 12-SEP-1996.
 PF 04-MAR-1996: US-02942.
 PR 19-MAY-1995: US-447172.
 PR 06-MAR-1995: US-399412.
 PR 19-MAY-1995: US-447169.
 PA (NEXS-) NEXSTAR PHARM INC.
 PI Gold L, Janjic N, Lochrie M, Parma D;
 DR WPI: 96-425373/42.
 PT New non-naturally occurring nucleic acid ligands to secretory
 phospholipase A2, VEGF and HIV-1 gag - useful e.g. for diagnosis
 and treatment of inflammation, angiogenesis and HIV infection.
 PS Claim 16; Page 71; 16pp; English.
 CC T88594-T88671 are non-naturally occurring nucleic acid ligands (NAL)
 for vascular endothelial growth factor (VEGF). Other non-naturally
 occurring ligands for secretory phospholipase A2 (sPLA2) and HIV-1 gag
 were also made. The NAL were identified using the SELEX procedure.
 CC They are used as diagnostic or treatment agents. For example, those for
 sPLA2 are used to study/diagnose inflammation and as anti-inflammatory
 agents e.g. for the treatment of rheumatoid arthritis and septic shock;
 CC those for VEGF are used to treat angiogenesis-mediated diseases e.g.
 psoriasis, haemangioma, metastasis; and those for gag are used to
 CC inhibit HIV-1 replication. NAL can also generally be used to identify
 CC the corresponding target.
 SQ Sequence 86 BP; 21 A; 22 C; 24 G; 18 T;

Query Match 74.1%; Score 12.6; DB 1; Length 86;
 Best Local Similarity 70.6%; Pred. No. 1.9e+02;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgytngcrtcnacytg 17
 ||||| ||||| |||||
 DB 50 TGGTTCGCTCCACTTG 66

RESULT 3
 ID 090000/c
 AC 090000 standard; DNA; 31 BP.
 DT 16-JAN-1996 (first entry)
 DE Modified Derf II major tick allergen V5A forward primer.
 KM Major tick allergen; Derf II; modified; mutein; anaphylactic shock;
 KM mutagenic polymerase chain reaction primer; ss.
 OS Synthetic.
 PN J07095887-A.
 PD 11-APR-1995.
 PF 29-SEP-1993: 275897.
 PR 29-SEP-1993: JP-275897.
 PA (ASAK) ASAKI BREWERIES LTD.
 PA (NIKK-) NIKKA WHISKEY KK.
 PA (TORI) TORII YAKUHTN KK.
 DR WPI: 95-174791/23.
 PT Modified Derf II tick allergen with Ala subst. at specific position
 PT - and its preparation by culturing transformant host cell, useful
 PT for treating anaphylactic shock
 PS Example 1: Page 5; 50pp; Japanese.
 CC DerfII major tick allergen sequences which are modified versions of
 CC three wild-type sequences are claimed (see R74567-R74569, along with
 CC their corresp. cDNA coding sequences Q89963-Q89965 for the unmodified
 CC allergens); the modified versions of these allergens have one
 CC specified amino acid residue replaced by an Ala residue and are useful
 CC for treating anaphylactic shock. See Q89966-Q89964 and R74570-R74598
 CC for the 29 specifically claimed examples of modified DerfII major tick
 CC allergens. The amino acid substitutions were introduced by mutation

CC of the corresp. codon in the wild-type coding sequence using PCR
 CC primers Q89996-Q90016.
 SQ Sequence 31 BP; 10 A; 4 C; 9 G; 8 T;

Query Match 71.8%; Score 12.2; DB 1; Length 31;
 Best Local Similarity 64.7%; Pred. No. 2.8e+02;
 Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgytngcrtcnacytg 17
 ||||| ||||| |||||
 DB 28 TCTTTAGCATCGACTTG 12

RESULT 4
 ID T84384/c
 AC T84384 standard; DNA; 98 BP.
 DT 12-NOV-1997 (first entry)
 DE Friedreich's ataxia STM7 gene exon 21.
 KM STM7 gene; Friedreich's ataxia; FRDA; neurodegeneration;
 KM phosphatidylinositol-4-phosphate 5-kinase; PtdInsPK;
 KM transgenic animal; animal model; diagnosis; therapy;
 KM STM7.IIIa, STM7.IIIB, STM7.IIIC; ss.
 OS Homo sapiens.
 PN WO9705234-A2.
 PD 13-FEB-1997.
 PF 24-JUL-1996: G01786.
 PR 28-JUL-1995: GB-015508.
 PR 26-JUL-1995: GB-015309.
 PA (UNLO) IMPERIAL COLLEGE SCI TECHNOLOGY & MED.
 PI Chamberlain S, Doudney CWF, Garcia-Valdecasas JJC;
 PI Hillermann R, Pook MA;
 DR WPI: 97-145674/13.
 PT New gene, STM7, and new protein causing Friedreich's Ataxia -
 PT encodes phosphatidylinositol phosphate kinase isoform; also splice
 PT variants and transgenic animals, for diagnosing pre-disposition to
 PT Friedreich's ataxia
 PS Claim 1; Fig 16; 82pp; English.
 CC DNA sequences (T84364-86) are provided for exons 1-23,
 CC respectively, of the human STM7 gene associated with Friedreich's
 CC ataxia (FRDA). The STM7.I gene (see T84358) comprises exons
 CC 1-16 of the gene, while splice variant STM7.IIIa includes exons
 CC 1-13, 17 and 19-22, splice variant STM7.IIIB includes exons
 CC 1-15, 17, 19, 21-22 and splice variant STM7.IIIC includes exons
 CC 1-15 and 19-22. STM7 nucleic acids can be used to produce STM7
 CC polypeptides (see W00978) in transfected host cells, to design
 CC probes (esp. oligonucleotides from exons 1-16, partic. 7-11) used
 CC in a claimed method for determining an inherited predisposition to
 CC FRDA, to generate transgenic animal models of FRDA and in somatic
 CC cell therapy.
 SQ Sequence 98 BP; 30 A; 23 C; 22 G; 23 T;

Query Match 71.8%; Score 12.2; DB 1; Length 98;
 Best Local Similarity 64.7%; Pred. No. 3.2e+02;
 Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgytngcrtcnacytg 17
 ||||| ||||| |||||
 DB 70 TGGTTCGCTCCACTTG 54

RESULT 5
 ID 090002/c
 AC 090002 standard; DNA; 34 BP.
 DT 16-JAN-1996 (first entry)
 DE Modified Derf II major tick allergen D7A forward primer.
 KM Major tick allergen; Derf II; modified; mutein; anaphylactic shock;
 KM mutagenic polymerase chain reaction primer; ss.
 OS Synthetic.
 PN J07095887-A.


```

PD 11-APR-1995.
PF 29-SEP-1993: 275897.
PA (ASAK ) ASahi BREWERIES LTD.
PA (NIKK-) NIKKA WHISKY KK.
PA (TORI ) TORII YAKUHIN KK.
DR WPI: 95-174791/23.
PT Modified Der fti tick allergen with Ala substid. at specific position
PT - and its preparation by culturing transformant host cell, useful
PT for treating anaphylactic shock
PS Example 1; Page 5; 50pp; Japanese.
CC Derfitt major tick allergen sequences which are modified versions of
CC three wild-type sequences are claimed (see R74567-R74569, along with
CC their corresp. cDNA coding sequences Q89963-Q89965 for the unmodified
CC allergens); the modified versions of these allergens have one
CC specified amino acid residue replaced by an Ala residue and are useful
CC for treating anaphylactic shock. See Q89966-Q89994 and R74570-R74598
CC for the 29 specifically claimed examples of modified Derfitt major tick
CC allergens. The amino acid substitutions were introduced by mutation
CC of the corresp. codon in the wild-type coding sequence using PCR
CC primers Q89996-Q90016.
SQ Sequence 34 BP; 9 A; 7 C; 9 G; 9 T;

Query Match 69.4%; Score 11.8; DB 1; Length 34;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 yttngcrtcnacyt 17
   ||| ||| ||| ||| |||
Db 26 TTTAGCATCTCTCT 12

RESULT 6
Q12936
ID 012936 standard; DNA: 50 BP.
AC 012936:
DT 11-OCT-1991 (first entry)
DE Analyte capture probe: HIV.97.XT1 for HIV.
KW Human Immunodeficiency Virus; ss.
OS Synthetic.
FH Key location/Qualifiers
FT misc-feature 1..30
FT /tag= a /note= "complementary to constant region of HIV "
FT /tag= b /note= "complementary to bound oligonucleotide"
FT misc-feature 31..50
FT /tag= b /note= "complementary to bound oligonucleotide"
PN MO9110746-A.
PD 25-JUL-1991.
PF 10-JAN-1991: U00213.
PR 10-JAN-1990: US-463022.
PA (CHIR-) CHIRON CORP.
PI Urdea MS:
DR WPI: 91-238034/32.
PT DNA-dependent RNA polymerase poly:deoxy:nucleotide construct -
PT used as signal amplifier in hybridisation assays, for detection
PT of N gonorrhoeae, HBV, HIV, etc.
PS Claim 69; Page 64; 88pp; English.
CC The analyte capture probe is used in a sandwich type hybridisation
CC assay in which a DNA-dependent RNA polymerase (T7) is used to
CC amplify the signal:
CC analyte capture template linker
CC ---w---|---x---|---a---
CC microtitre |---w---|---x---|---a---|---b---|---c---
CC dish } analyte template probe
CC The portion marked "y" hybridises with a specific portion of the
CC HIV genome and portion "w" hybridises with an oligonucleotide bound
CC to the microtitre dish. A template linker connects the analyte
CC with a T7 template probe which incorporates a T7 promoter and allows
CC amplification of the signal.
CC See also Q12914-Q12993.
SQ Sequence 50 BP; 8 A; 11 C; 13 G; 18 T;

```

```

Query Match 65.9%; Score 11.2; DB 1; Length 50;
Best Local Similarity 62.5%; Pred. No. 1.1e+03;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tytgngcrtcnacyt 16
   ||| ||| ||| ||| |||
Db 21 TGTTCAGCATCTCTCT 36

RESULT 7
T74386
ID T74386 standard; DNA: 50 BP.
AC T74386:
DT 26-AUG-1997 (first entry)
DE HIV analyte capture probe: HIV.97.2.XT1.
KW Promoter; template probe; signal amplifier; hybridisation assay; detect;
KW functional domain; DNA-dependent RNA polymerase; quantifying; analyte;
KW ligand receptor; amplification; hepatitis B virus; Neisseria gonorrhoeae;
KW bacterial beta-lactamase TEM-1 gene; Chlamydia; HIV; hepatitis C virus;
KW bacterial tet M determinant; consensus; T7 promoter; ss.
OS Synthetic.
PN US5629153-A.
PD 13-MAY-1997.
PF 10-JAN-1990: 463022.
PR 10-JAN-1991: US-463022.
PR 10-JAN-1991: US-639560.
PR 08-MAR-1994: US-207901.
PA (CHIR ) CHIRON CORP.
PI Urdea MS:
DR WPI: 97-280266/25.
PT DNA construct for use as signal amplifier in hybridisation assays -
PT containing DNA-dependent RNA polymerase promoter and template
PT sequences
PS Example 9; Column 41-42; 45pp; English.
CC A novel DNA construct (referred to as a "template probe") for use as a
CC signal amplifier in hybridisation assays to detect a target comprises 3
CC functional domains (A, B and C) orientated A-B-C or B-C-A. (A) is single-
CC stranded and is designed to hybridise to complementary target sequence.
CC (B) is double-stranded and functions as a DNA-dependent RNA polymerase
CC promoter. (C) is single- or double-stranded, and functions as a template
CC for the promoter activity of domain B. It consists of a nucleotide
CC sequence not found in the template. The DNA construct is used in a method
CC for detecting and quantifying an oligonucleotide analyte or a ligand
CC receptor by amplification of a biological signal in a nucleic acid
CC hybridisation assay. The method is especially useful for determination of
CC nucleic acid segments characteristic of hepatitis B virus, Neisseria
CC gonorrhoeae, bacterial beta-lactamase TEM-1 gene, Chlamydia, bacterial
CC tet M determinant, HIV or hepatitis C virus. Comparison of template
CC probes with various numbers of base pairs between the A and B domain was
CC carried out by assaying for the presence of Human Immunodeficiency Virus
CC (HIV) DNA in human plasma. T74384-95 are specific analyte capture probes
CC used in the assay.
SQ Sequence 50 BP; 8 A; 11 C; 13 G; 18 T;

Query Match 65.9%; Score 11.2; DB 1; Length 50;
Best Local Similarity 62.5%; Pred. No. 1.1e+03;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tytgngcrtcnacyt 16
   ||| ||| ||| ||| |||
Db 21 TGTTCAGCATCTCTCT 36

RESULT 8
Q97417
ID Q97417 standard; DNA: 54 BP.
AC Q97417:
DT 01-APR-1996 (first entry)
DE Fragment of wild type lacZ alpha gene.
KW lacZ alpha gene; recombination; chimERIC sequence; Interleukin-1;

```

KM antibody; tPA; growth hormone; repetitive DNA; ribozyme; aptamer;
KW gene therapy; ss.
OS Synthetic.
PN WO9522625-A1.
PD 24-AUG-1995.
PF 17-FEB-1995; U02126.
PR 17-FEB-1994; US-198431.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
PI Cramerl A, Stemmer WPC;
DR WPI: 95-302727/39.
PT DNA mutagenesis via random fragmentation and reassembly - useful in
PS the prodn. of mutant proteins having enhanced biological activity
Example 2: Fig 3, 120pp; English.
CC The sequences shown in Q97417 and Q97418 are fragments of the LacZ alpha
CC gene. Crossover between these two sequences was measured and indicated
CC that gene shuffling was occurring. The sequence was reinserted into
CC pUC18, and E.coli transformed with this plasmid were shown to have a
CC lower frequency of LacZ resistance than by mutagenesis without shuffling.
CC This method can also be used to shuffle whole gene sequences. The method
CC can also be used to produce recombinant proteins, and chimeric DNA
CC sequences. The methods can be used to produce proteins with a desired
CC phenotype having an advantageous predetermined selectable property.
CC Proteins such as interleukin-1, antibodies, tPA and growth hormones can
CC be generated that have altered specificity or activity. Mutant sequences
CC such as promoter regions, introns, exons and enhancer sequences can also
CC be generated, which can be used to generate genes having increased
CC expression rates. The method can be used in the study of repetitive DNA
CC regions, and for the mutation of ribozymes or aptamers. Shuffled viral
CC proteins which may comprise epitopes that will arise by viral evolution
CC can also be produced. This method of shuffling sequences may be used for
CC generating gene therapy vectors and replication-defective gene therapy
CC constructs. The advantage with this method is that it is not necessary
CC to know the actual DNA or RNA sequence of the specific nucleic acid
SQ Sequence 54 BP; 9 A; 18 C; 15 G; 12 T;

Query Match 65.9%; Score 11.2; DB 1; Length 54;
Best Local Similarity 62.5%; Pred. No. 1.1e+03;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgyttingcrtcnacyt 16
Db 7 TGCTGCGCATCCACT 22

RESULT 9
V54949
ID V54949 standard; DNA; 54 BP.
AC V54949;
DT 19-NOV-1998 (first entry)
DE LacZ alpha gene stop codon mutant.
KW LacZ alpha gene; recombination; gene shuffling; ss.
OS Synthetic.
PN US5811238-A.
PD 22-SEP-1998.
PF 30-NOV-1995; 564955.
PR 04-MAR-1996; US-537874.
PR 17-FEB-1994; US-198431.
PR 30-NOV-1995; US-564955.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
PI Cramerl A, Stemmer WPC;
DR WPI: 98-530860/45.
PT Production of polynucleotides with desired properties - by iterative
PS selection and recombination
CC Disclosure; Fig 3b; 74pp; English.
CC V54947-50 represent fragments of LacZ alpha gene stop codon mutants,
CC produced using the method of the invention. The specification
CC describes a method for evolving a polynucleotide for acquisition
CC of a desired property. The method comprises providing a population
CC of variants of the polynucleotide, at least one of which is in
CC cell-free form, shuffling the variants of the polynucleotide to form
CC recombinant polynucleotides, selecting or screening for recombinant

CC polynucleotides that have evolved toward the desired property and
CC repeating the steps with the selected recombinant polynucleotides
CC until a recombinant polynucleotide has acquired the desired property.
CC The recombinant polynucleotide produced by method, or its expression
CC product, can be formulated as a pharmaceutical.
SQ Sequence 54 BP; 9 A; 18 C; 15 G; 12 T;

Query Match 65.9%; Score 11.2; DB 1; Length 54;
Best Local Similarity 62.5%; Pred. No. 1.1e+03;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgyttingcrtcnacyt 16
Db 7 TGCTGCGCATCCACT 22

RESULT 10
Q97418
ID Q97418 standard; DNA; 57 BP.
AC Q97418;
DT 01-APR-1996 (first entry)
DE Fragment of mutated lacZ alpha gene.
KW LacZ alpha gene; recombination; chimeric sequence; interleukin-1;
KW antibody; tPA; growth hormone; repetitive DNA; ribozyme; aptamer;
KW gene therapy; ss.
OS Synthetic.
PN WO9522625-A1.
PD 24-AUG-1995.
PF 17-FEB-1995; U02126.
PR 17-FEB-1994; US-198431.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
PI Cramerl A, Stemmer WPC;
DR WPI: 95-302727/39.
PT DNA mutagenesis via random fragmentation and reassembly - useful in
PS the prodn. of mutant proteins having enhanced biological activity
Example 2: Fig 3, 120pp; English.
CC The sequences shown in Q97417 and Q97418 are fragments of the LacZ alpha
CC gene. Crossover between these two sequences was measured and indicated
CC that gene shuffling was occurring. The sequence was reinserted into
CC pUC18, and E.coli transformed with this plasmid were shown to have a
CC lower frequency of LacZ resistance than by mutagenesis without shuffling.
CC This method can also be used to shuffle whole gene sequences. The method
CC can also be used to produce recombinant proteins, and chimeric DNA
CC sequences. The methods can be used to produce proteins with a desired
CC phenotype having an advantageous predetermined selectable property.
CC Proteins such as interleukin-1, antibodies, tPA and growth hormones can
CC be generated that have altered specificity or activity. Mutant sequences
CC such as promoter regions, introns, exons and enhancer sequences can also
CC be generated, which can be used to generate genes having increased
CC expression rates. The method can be used in the study of repetitive DNA
CC regions, and for the mutation of ribozymes or aptamers. Shuffled viral
CC proteins which may comprise epitopes that will arise by viral evolution
CC can also be produced. This method of shuffling sequences may be used for
CC generating gene therapy vectors and replication-defective gene therapy
CC constructs. The advantage with this method is that it is not necessary
CC to know the actual DNA or RNA sequence of the specific nucleic acid
SQ Sequence 57 BP; 13 A; 17 C; 10 G; 17 T;

Query Match 65.9%; Score 11.2; DB 1; Length 57;
Best Local Similarity 62.5%; Pred. No. 1.1e+03;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgyttingcrtcnacyt 16
Db 7 TGCTGCGCATCCACT 22

RESULT 11
V54950
ID V54950 standard; DNA; 57 BP.

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:52:56 ; Search time 4521.53 Seconds
(without alignments)
15.239 Million cell updates/sec

Title: US-09-164-714-3

Perfect score: 17
Sequence: 1 tgyttingcrtcnacytg 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 08
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *
44: gb_est25: *

45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_gss1: *
83: gb_gss2: *
84: gb_gss3: *
85: gb_gss4: *
86: em_gss1: *
87: em_gss2: *
88: em_gss3: *
89: em_gss4: *
90: gb_gss5: *
91: gb_gss6: *
92: gb_gss7: *
93: gb_gss8: *
94: gb_gss9: *
95: em_gss5: *
96: em_gss6: *
97: em_gss7: *
98: em_gss8: *
99: em_gss9: *
100: em_gss10: *
101: em_gss11: *
102: gb_gss10: *
103: gb_gss11: *
104: em_gss12: *
105: gb_gss12: *
106: gb_gss13: *
107: gb_gss14: *
108: gb_gss15: *
109: gb_gss16: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result

Query

No.	Score	Match Length	DB	ID	Description
1	12.8	75.3	80	21	T83169
2	12.8	75.3	80	27	AA045928
3	12.2	71.8	61	23	H23812
4	12.2	71.8	82	23	H40781
5	11.8	69.4	80	84	B14303
6	11.8	69.4	92	41	A1054253
7	11.8	69.4	97	50	A1669737
8	11.6	68.2	61	34	AA482394
9	11.6	68.2	67	42	A1159297
10	11.2	65.9	63	45	A1338554
11	11.2	65.9	65	33	AA420410
12	11.2	65.9	94	63	A19398403
13	11.2	65.9	100	36	AA659018
14	11.2	65.9	55	40	AA911593
15	11	64.7	73	39	AA839154
16	11	64.7	79	40	AA922367
17	11	64.7	81	69	AA119807
18	11	64.7	100	94	AQ025853
19	10.8	63.5	72	24	U38119
20	10.6	62.4	34	34	AA481226
21	10.6	62.4	53	34	D18203
22	10.6	62.4	61	39	AA906850
23	10.6	62.4	70	22	H06004
24	10.6	62.4	70	43	A1216035
25	10.6	62.4	73	41	A1001677
26	10.6	62.4	74	35	AA578173
27	10.6	62.4	78	23	H49037
28	10.6	62.4	79	27	AA009424
29	10.6	62.4	79	63	A1971513
30	10.6	62.4	79	63	A1973702
31	10.6	62.4	83	20	T50051
32	10.6	62.4	86	30	AA244290
33	10.6	62.4	88	62	A18773394
34	10.6	62.4	92	24	H88263
35	10.6	62.4	93	43	A1242111
36	10.6	62.4	94	34	AA487066
37	10.6	62.4	96	28	AA036143
38	10.6	62.4	99	20	T48764
39	10.6	62.4	99	38	AA745750
40	10.6	62.4	100	20	T51146
41	10.6	62.4	100	22	R23019
42	10.4	61.2	53	21	T58978
43	10.4	61.2	88	82	A1149323
44	10.2	60.0	41	82	HSK19903
45	10.2	60.0	46	40	AA953796

ALIGNMENTS

```

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel.: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Insert Size: 935
Source: IMAGE Consortium, LNL This clone is available royalty-free
through LNL; contact the IMAGE Consortium (info@image.lnl.gov)
for further information. Putative full length read
Insert Length: 935 SCD Error: 0.00
Seq primer: M13RPI
High quality sequence stop: 251.
Location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="GDB:466410"
/db_xref="taxon:9606"
/cclone="IMAGE:110793"
/clone_id="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pTRF3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
15' AACGTGACAGATTAATAAATGAAGATCTTTTTCCTTTTTTTT 3'1,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pTRF3 vector. Library
went through one round of normalization. library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      23 a          19 c          11 g          27 t
ORIGIN
Query Match     75.3%; Score 12.8; DB 21; Length 80;
Best Local Similarity 68.8%; Pred. NO. 1.1e+03;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
OY      1 tgytngcrtcnacyt 16
||:|| ||:|| ||:||
DB      4 TGTATTGCATCAACTT 19
RESULT      2
AA045928      80 bp      mRNA      EST      11-MAY-1997
LOCUS

```

RESULT	1
LOCUS	T83169
DEFINITION	T83169 80 bp mRNA EST 16-MAR-1995
ACCESSION	YD4105.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
VERSION	IMAGE:110793 5', mRNA sequence.
KEYWORDS	T83169 GI:711457
SOURCE	EST.
ORGANISM	human. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 80) Hallier,L., Clark,N., Dubaque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE	The Wasnu-Werc EST Project
JOURNAL	Unpublished (1995)
COMMENT	Contact: Wilson RK

SOURCE ORGANISM	human.
REFERENCE	Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 80) Hallier,L., Lennon,G., Becker,M., Donald,M.F., Chippelli,B., Chisose,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevaskis,E., Underwood,K., Wohlmann,P., Waterson,R., Wilson,R. and Maria,M.
TITLE	Generation and analysis of 280,000 human expressed sequence tags
JOURNAL	Genome Res. 6 (9), 807-828 (1996)
MEDLINE	97044478
COMMENT	On Apr 14, 1993 this sequence version replaced gi:716787. Contact: Wilson RK Washington State University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810

Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Putative full length read
 The vector to vector length is 81
 Insert Length: 822 Std Error: 0.00
 Seq primer: -28M13 rev2 from Amersham.

FEATURES

source

Location/Qualifiers
 1..80
 /organism="Homo sapiens"
 /db_xref="GDB:3803227"
 /db_xref="taxon:9606"
 /clone="IMAGE:488640"
 /clone_lib="Soares_pregnant_uterus_NBHPU"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: pT73-Pac; Site_1: Not I;
 Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5',
 AACTGAGAGATTGCGCGCCCTTTTCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by M. Fatima Bonaldo."

BASE COUNT

23 a 19 c 11 g 27 t

ORIGIN

Query Match

Best Local Similarity 75.3%; Score 12.8; DB 27; Length 80;

Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tgytngcrtcnacyt 16

Db 4 TGTTCGATCGACTT 19

RESULT 3

H23812

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

H23812 61 bp mRNA EST 06-JUL-1995
 Y070b11.s1 Soares adult brain N25HB55Y Homo sapiens cDNA clone
 IMAGE:173829 3' similar to SP:K02F3.2 CE01348 ; mRNA sequence.
 H23812
 H23812.1 GI:892507
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 61)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
 Wilson, R.
 The WashU-Merck EST Project
 Unpublished (1995)
 On Sep 21, 1992 this sequence version replaced gi:276165.
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 1692
 High quality sequence starts: 1
 High quality sequence stops: 1
 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand
 Insert Length: 1692 Std Error: 0.00
 Seq primer: Promega -21m13
 High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers
 1..61
 /organism="Homo sapiens"
 /db_xref="GDB:3835864"
 /db_xref="taxon:9606"
 /clone="IMAGE:173829"
 /clone_lib="Soares adult brain N25HB55Y"
 /sex="Male"
 /dev_stage="55-year old"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: brain; Vector: pT73D (Pharmacia) with a
 modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTTCGATCGATGAGTGGAGCGCGCCCTTTTCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 53. Library constructed by Bento
 Soares and M. Fatima Bonaldo. The adult brain RNA was
 provided by Dr. Donald H. Gilden. Tissue was acquired
 17-18 hours after death which occurred in consequence of a
 ruptured aortic aneurysm. RNA was prepared from a pool of
 tissues representing the following areas of the brain:
 frontal, parietal, temporal and occipital cortex from the
 left and right hemispheres, subcortical white matter,
 basal ganglia, thalamus, cerebellum, midbrain, pons and
 medulla."

BASE COUNT 14 a 21 c 16 g 10 t

ORIGIN

Query Match

Best Local Similarity 71.8%; Score 12.2; DB 23; Length 61;

Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 tgytngcrtcnacytg 17

Db 12 TGTTCGATCGACTG 28

RESULT 4

H40781

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

H40781 82 bp mRNA EST 31-JUL-1995
 Y008C10.s1 Soares adult brain N25HB55Y Homo sapiens cDNA clone
 IMAGE:177330 3' similar to SP:K02F3.2 CE01348 ; mRNA sequence.
 H40781
 H40781.1 GI:916833
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 82)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
 Wilson, R.
 The WashU-Merck EST Project
 Unpublished (1995)
 On May 9, 1995 this sequence version replaced gi:803137.
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 1926

High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Insert length: 1926 Std Error: 0.00
Seq primer: Promega -21ml3

FEATURES

source

Location/Qualifiers

1. 82

/organism="Homo sapiens"
/db_xref="GDB:3839526"
/db_xref="taxon:9606"
/clone="IMAGE:177330"
/clone_lib="Soares adult brain N2b5HB5Y"
/sex="Male"
/dev_stage="55-year old"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: brain; Vector: pT73D (Pharmacia) with a
modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TCTTACCATCTGAAGTGGACGGCGCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 53. Library constructed by Bento
Soares and M. Fatima Bonaldo. The adult brain RNA was
provided by Dr. Donald H. Gilden. Tissue was acquired
17-18 hours after death which occurred in consequence of a
ruptured aortic aneurysm. RNA was prepared from a pool of
tissues representing the following areas of the brain:
frontal, parietal, temporal and occipital cortex from the
left and right hemispheres, subcortical white matter,
basal ganglia, thalamus, cerebellum, midbrain, pons and
medulla."

BASE COUNT 17 a 28 c 23 g 13 t 1 others
ORIGIN

Query Match 71.8%; Score 12.2; DB 23; Length 82;
Best Local Similarity 64.7%; Pred. No. 2.5e+03;

Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacyt 17
11:11 11:11 11:11
Db 33 TGTTCCATCGACCTG 49

RESULT 5
LOCUS B14303 80 bp DNA GSS 04-JUN-1998
DEFINITION A-345G4.TV C19978SKA1 Homo sapiens genomic clone A-345G4, genomic
survey sequence.
ACCESSION B14303
VERSION B14303.1 GI:2122052
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 80)

REFERENCE 1
AUTHORS Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
TITLE Use of a BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other GSSs: 345G4.TP
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/cdb/humgen/Bac_end_search/Bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1. 80

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="A-345G4"
/clone_lib="C1978SKA1"
/sex="Female"
/cell_type="Fibroblast"
/note="Vector: pBAC108L; Site:1: HindIII; Site:2: HindIII;
Caltech Human BAC Library A1"
BASE COUNT 21 a 21 c 23 t 3 others
ORIGIN

Query Match 69.4%; Score 11.8; DB 84; Length 80;
Best Local Similarity 62.5%; Pred. No. 4.3e+03;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacyt 16
11:11 11:11 11:11
Db 12 TGCTTCGACNCCACT 27

RESULT 6
LOCUS A1054253 92 bp mRNA EST 27-JUL-1998
DEFINITION G175H05.x1 NCI_CGAP_Ov26 Homo sapiens cDNA clone IMAGE:1862361 3',
mRNA sequence.
ACCESSION A1054253
VERSION A1054253.1 GI:3322040
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 92)

REFERENCE 1
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced g1:2287056.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Monica Brown, M.D., Elise Kohn, M.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Kitzman, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbdp/image/image.html

Insert Length: 184 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.
Location/Qualifiers

FEATURES

source

1. 92

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1862361"
/clone_lib="NCI_CGAP_Ov26"
/sex="female"
/tissue_type="papillary serous carcinoma"
/dev_stage="adult"
/lab_host="DH10B"

/note="Organ: ovary; Vector: PAMPI; mRNA made from papillary serous ovarian carcinoma, cDNA made by oligo-dt priming. Directionally cloned. Size-selected on agarose gel, average insert size 500 bp. Primary library, non-amplified."

BASE COUNT 9 a 20 c 14 g 49 t

ORIGIN

Query Match 69.4%; Score 11.8; DB 41; Length 92;
Best Local Similarity 66.7%; Pred. No. 4.4e+03;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacy 15
|||||
Db 23 TGCTTGCATCCACT 37

RESULT 7
AI669737/c 97 bp mRNA EST 14-MAY-1999
LOCUS AI669737
DEFINITION tui2608.x1 NCI-CGAP-Pr28 Homo sapiens cDNA clone IMAGE:2250854 3'
similar to gb:U02947 EXTRACELLULAR SUPEROXIDE DISMUTASE PRECURSOR (HUMAN); mRNA sequence.
AI669737
VERSION AI669737.1 GI:4834511
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 97)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
On Dec 20, 1995 this sequence version replaced gi:1131205.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/dbfp/image/image.html

Seq primer: -400p from Gibco.
Location/Qualifiers
1. 97

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2250854"
/clone_lib="NCI-CGAP-Pr28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI-CGAP-Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 985608-986759, 110192-1101959, and 121928-1220615). Subtraction by Bento Soares and M. Fatima Bernaldo."

BASE COUNT 29 a 20 c 26 g 22 t

ORIGIN

Query Match

69.4%; Score 11.8; DB 50; Length 97;

Best Local Similarity 66.7%; Pred. No. 4.5e+03;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacy 15
|||||
Db 91 TGTTTCGATCCACC 77

RESULT 8
AA482394/c 61 bp mRNA EST 09-NOV-1997
LOCUS AA482394
DEFINITION r534a02.r1 Soares ovary tumor NBHOF Homo sapiens cDNA clone
IMAGE:724202.5' similar to SR:N18M_BOVIN Q02370 NADH-UBIQUINONE
OXIDOREDUCTASE B8 SUBUNIT ; mRNA sequence.
AA482394
VERSION AA482394.1 GI:2210072
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 61)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Seppoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1395310.
COMMENT Contact: Wilson R.
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Insert length: 421 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 1.

FEATURES
source
Location/Qualifiers
1. 61

/organism="Homo sapiens"
/db_xref="GDB:5935704"
/db_xref="taxon:9606"
/clone="IMAGE:724202"
/clone_lib="Soares ovary tumor NBHOF"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: ovary; Vector: p773D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCGGTTTGTGTGTGTGTGT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p773 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bernaldo."

BASE COUNT 12 a 21 c 13 g 15 t

ORIGIN

Query Match 68.2%; Score 11.6; DB 34; Length 61;
Best Local Similarity 68.8%; Pred. No. 5.2e+03;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 gytngcrtcnacy 17
|||||
Db 23 GATTGATCAACTTG 8

RESULT 9
LOCUS A1159297/c
DEFINITION A1159297 67 bp mRNA EST 02-OCT-1998
vz85e01.r1 Soares mammary.gland_NBMG Mus musculus cDNA clone
IMAGE:1333272.5' similar to WP:MI53.1 C603511
PYROLINE-5-CARBOXYLATE REDUCTASE ;, mRNA sequence.

ACCESSION A1159297
VERSION A1159297.1 GI:3692479
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 67)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2287146.
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:692816
Trace considered overall poor quality
Seq primer: -28m13 rev2.ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..67
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1333272"
/clone_1lb="Soares_mammary.gland_NBMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTTCACATCTGAGTGGAGCGCGCGATGGTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

BASE COUNT 21 a 19 c 18 g 9 t
ORIGIN

Query Match 68.2%; Score 11.6; DB 42; Length 67;
Best Local Similarity 68.8%; Pred. No. 5.4e+03;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgytngcrtcnacyt 16
|||||
DB 48 TGTTCAGCTTCTACTT 33

RESULT 10

A138554/c
LOCUS A138554 63 bp mRNA EST 29-DEC-1998
DEFINITION q93c08.x1 Soares total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
IMAGE:1938926.3' similar to gb:X02067.H.sapiens mRNA for 7SL RNA
pseudogene (HUMAN); contains element MER30 repetitive element ;,
mRNA sequence.

ACCESSION A138554
VERSION A138554.1 GI:4075481
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 63)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Feb 11, 1998 this sequence version replaced gi:2873097.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Trace considered overall poor quality
Seq primer: -40Up from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1..63
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1938926"
/clone_1lb="Soares_total_fetus_Nb2HF8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAGTGGAGCGCGCGCTTATTTTTTTTTTTTTTTTTTTT
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 18 a 21 c 12 g 12 t
ORIGIN

Query Match 65.9%; Score 11.2; DB 45; Length 63;
Best Local Similarity 62.5%; Pred. No. 9.2e+03;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 gytngcrtcnacyt 17
|||||
DB 49 GTTTCACATCAATTG 34

RESULT 11
LOCUS AA420410/c
DEFINITION AA420410 65 bp mRNA EST 16-OCT-1997
vcs1e07.r1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
IMAGE:778116.3', mRNA sequence.

ACCESSION AA420410
VERSION AA420410.1 GI:2080907
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 65)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 376 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES

source

```
1..100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1216937"
/clone_lib="NCI-CGAP_Alvi1"
/tissue_type="alveolar rhabdomyosarcoma"
/lab_host="DH10B"
/note="Vector: PAMPI0; mRNA made from alveolar
rhabdomyosarcoma, cDNA made by oligo-dT priming.
Non-directionally cloned. Size-selected on agarose gel,
average insert size 600 bp. Reference: Krizman et al.
(1996) Cancer Research 56:5380-5383."
BASE COUNT      21 a      24 c      20 g      35 t
ORIGIN
```

Query Match

Best Local Similarity 71.4%; Pred. No.1e+04; Length 100;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 ttngctcnaacytg 17
||| ||| ||| |||
Db 12 TTCGATCCACCTG 25

RESULT 14

AA911593

LOCUS AA911593 55 bp mRNA EST 21-APR-1998
DEFINITION o490c01.s1 NCI-CGAP_Br5 Homo sapiens cDNA clone IMAGE:1377504
Similar to TR:05157 055157 HYPOTHETICAL 23.6 KD PROTEIN. ; contains
MER22.t3 MSRI repetitive element ;, mRNA sequence.

ACCESSION AA911593
VERSION AA911593.1 GI:3050957
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 55)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

REFERENCE 1
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE Unpublished (1997)
JOURNAL On Jan 19, 1998 this sequence version replaced gi:2285536.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

JOURNAL

COMMENT

Email: Robert.Strausberg@nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Kristina A. Cole, M.D.,
Ph.D. student, Rodrigo F. Chuaguí, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 441 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES

Location/Qualifiers

source

```
1..55
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1377504"
/clone_lib="NCI-CGAP_Br5"
/sex="female"
/tissue_type="infiltrating ductal carcinoma"
/lab_host="DH10B"
/note="Organ: breast; Vector: PAMPI0; mRNA made from
infiltrating ductal carcinoma, cDNA made by oligo-dT
priming. Non-directionally cloned. Size-selected on
agarose gel, average insert size 600 bp."
BASE COUNT      9 a      22 c      12 g      12 t
ORIGIN
```

Query Match

Best Local Similarity 64.7%; Pred. No.1.2e+04; Length 55;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 tgytngctcnaacytg 17
||| ||| ||| |||
Db 24 TGTCTGATCAACATG 40

RESULT 15

AA839154

LOCUS AA839154 73 bp mRNA EST 27-FEB-1998
DEFINITION v448c07.r1 Soares_mammary_gland_NBMGC Mus musculus cDNA clone
IMAGE:1247052 5' similar to SW:STO_HUMAN P47897 GLUTAMINYL-TRNA
SYNTHETASE ;, mRNA sequence.

ACCESSION AA839154
VERSION AA839154.1 GI:2915249
KEYWORDS EST.

SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 73)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Weisling,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)

COMMENT On Jan 19, 1998 this sequence version replaced gi:2150179.
Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:660740

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.

FEATURES

source

```
1..73
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1247052"
/clone_lib="Soares_mammary_gland_NBMGC"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT73D-Pac
```

(Pharmacia) with a modified polylinker; Site 1: Not I;
 Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5',
 TCTTTCACATCTGAAGTGGAGCGCCGCGCAATGGTTTTTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT73 vector.
 RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
 constructed and normalized by Bento Soares and M. Fatima
 Bonaldo.

BASE COUNT	17	a	20	c	15	g	21	t
ORIGIN								

Query Match 64.7%; Score 11; DB 39; Length 73;
 Best Local Similarity 64.7%; Pred. No. 1.3e+04;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY	1	tggttgctcncactg	17
		: :	
Db	28	TCTTTGCTTCACCTG	44

Search completed: June 4, 2000, 13:52:59
 Job time: 20923 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:37 ; Search time 244.64 Seconds
(without alignments)
9.033 Million cell updates/sec

Title: US-09-164-714-3

Perfect score: 17

Sequence: 1 tgytngctcnaacytg 17

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: Issued_Patents_NA:*
2: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5C.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/5D.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/6.COMB.seq:*
7: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12.6	74.1	86	1	US-08-479-783A-10
2	12.6	74.1	86	1	US-08-479-725-10
3	12.6	74.1	86	2	US-08-618-693-10
4	12.6	74.1	86	2	US-08-447-169A-168
5	12.6	74.1	86	6	PCT-US96-08014-99
6	11.2	65.9	27	4	US-08-859-998-1003
7	11.2	65.9	50	1	US-08-207-901-50
8	11.2	65.9	53	3	US-08-454-557C-74
9	11.2	65.9	53	4	US-08-340-426D-74
10	11.2	65.9	53	4	US-08-450-673C-74
11	11.2	65.9	53	6	PCT-US95-17111A-74
12	11.2	65.9	54	2	US-08-564-955-60
13	11.2	65.9	57	3	US-08-537-874-58
14	11.2	65.9	57	3	US-08-537-874-58
15	11.2	65.9	57	3	US-08-537-874-58
16	11.2	65.9	57	3	US-08-537-874-58
17	10.8	63.5	21	4	US-09-028-361A-7
18	10.6	62.4	22	1	US-08-379-078-430
19	10.6	62.4	22	1	US-08-379-078-431
20	10.6	62.4	22	1	US-08-379-078-432
21	10.6	62.4	27	1	US-07-766-351-9
22	10.6	62.4	27	1	US-08-059-032-9
23	10.6	62.4	27	6	PCT-US91-07290-9
24	10.6	62.4	30	2	US-08-480-784-80
25	10.6	62.4	30	2	US-08-483-553-80
26	10.6	62.4	30	2	US-08-487-002-80
27	10.6	62.4	30	2	US-08-483-554B-80

28	10.6	62.4	30	2	US-08-488-011B-80	Sequence 80, Appl
29	10.6	62.4	30	6	PCT-US95-10202-80	Sequence 80, Appl
30	10.6	62.4	30	6	PCT-US95-10203-80	Sequence 80, Appl
31	10.6	62.4	30	6	PCT-US95-10220-80	Sequence 80, Appl
32	10.6	62.4	36	1	US-07-747-785-12	Sequence 12, Appl
33	10.6	62.4	42	1	US-07-885-689A-6	Sequence 6, Appl
34	10.6	62.4	45	4	US-08-495-695B-15	Sequence 15, Appl
35	10.6	62.4	45	6	PCT-US94-14436-15	Sequence 15, Appl
36	10.6	62.4	56	1	US-07-854-845B-3	Sequence 3, Appl
37	10.6	62.4	75	1	US-07-766-351-13	Sequence 13, Appl
38	10.6	62.4	75	1	US-08-059-032-13	Sequence 13, Appl
39	10.6	62.4	75	6	PCT-US91-07290-13	Sequence 13, Appl
40	10.6	62.4	87	2	US-08-474-140-13	Sequence 13, Appl
41	10.6	62.4	87	2	US-08-477-630-13	Sequence 13, Appl
42	10.6	62.4	87	2	US-08-472-293-13	Sequence 13, Appl
43	10.6	62.4	87	2	US-08-474-545-13	Sequence 13, Appl
44	10.6	62.4	87	2	US-08-478-341-13	Sequence 13, Appl
45	10.6	62.4	90	1	US-08-123-702-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-08-479-783A-10
Sequence 10, Application US/08479783A
Patent No. 5668264
GENERAL INFORMATION:
APPLICANT: NEBOJSA JANJIC
APPLICANT: LARRY GOLD
TITLE OF INVENTION: HIGH AFFINITY PDGF NUCLEIC
TITLE OF INVENTION: ACID LIGANDS
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Swanson and Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,783A
FILING DATE: 7-JUNE-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/931,473
FILING DATE: 17-AUGUST-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,991
FILING DATE: 8-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
ATTORNEY/AGENT INFORMATION:
NAME: Diane H. McLearn
REGISTRATION NUMBER: 33,960
REFERENCE/DOCKET NUMBER: NEX42-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:

LENGTH: 86 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-479-783A-10

Query Match 74.1%; Score 12.6; DB 1; Length 86;
Best Local Similarity 70.6%; Pred. No. 1.1e+02;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 tgytngcrtcnacytg 17
||| ||: || ||: ||
Db 50 TGGTAGCGTCACCTTG 66

RESULT 2

US-08-479-725-10
; Sequence 10, Application US/08479725
; Patent No. 5674685
; GENERAL INFORMATION:
; APPLICANT: NEBOJSA JANJIC
; APPLICANT: LARRY GOLD
; TITLE OF INVENTION: HIGH AFFINITY PDGF NUCLEIC
; TITLE OF INVENTION: ACID LIGANDS
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue, Suite #200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,725
; FILING DATE: 7-JUNE-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/931,473
; FILING DATE: 17-AUGUST-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/117,991
; FILING DATE: 8-SEPTEMBER-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Diane H. McClearn
; REGISTRATION NUMBER: 33,960
; REFERENCE/DOCKET NUMBER: NEX42-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-479-725-10

Query Match 74.1%; Score 12.6; DB 1; Length 86;
Best Local Similarity 70.6%; Pred. No. 1.1e+02;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 tgytngcrtcnacytg 17
||| ||: || ||: ||
Db 50 TGGTAGCGTCACCTTG 66

RESULT 3

US-08-618-693-10
; Sequence 10, Application US/08618693
; Patent No. 5723594
; GENERAL INFORMATION:
; APPLICANT: NEBOJSA JANJIC
; APPLICANT: LARRY GOLD
; TITLE OF INVENTION: HIGH AFFINITY PDGF NUCLEIC
; TITLE OF INVENTION: ACID LIGANDS
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue, Suite #200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,693
; FILING DATE: 20 MARCH 1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,783
; FILING DATE: 7-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX42/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-618-693-10

Query Match 74.1%; Score 12.6; DB 2; Length 86;
Best Local Similarity 70.6%; Pred. No. 1.1e+02;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 tgytngcrtcnacytg 17
||| ||: || ||: ||
Db 50 TGGTAGCGTCACCTTG 66

RESULT 4
US-08-447-169A-168
; Sequence 168, Application US/08447169A
; Patent No. 5811533
; GENERAL INFORMATION:
; APPLICANT: JANJIC, N. and GOLD, L.
; TITLE OF INVENTION: HIGH-AFFINITY OLIGONUCLEOTIDE

ADDRESSEE: Fish & Richardson, P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: CA
COUNTRY: US
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,998
FILING DATE: 21-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E.
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 09096/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 1003:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: oligonucleotide primer
US-08-859-998-1003

Query Match 65.9%; Score 11.2; DB 4; Length 27;
Best Local Similarity 62.5%; Pred. No. 6.3e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgyttgcttcnacyt 16
||:| | | :| | | :| |
DB 17 TGCTTTCATCCACCT 2

RESULT 7
US-08-207-901-50
Sequence 50, Application US/08207901
Patent No. 5629153
GENERAL INFORMATION:
APPLICANT: Urdca, Michael S.
TITLE OF INVENTION: USE OF DNA-DEPENDENT RNA POLYMERASE
TITLE OF INVENTION: TRANSCRIPTS AS REPORTER MOLECULES FOR SIGNAL
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/207,901
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/639,560B
FILING DATE: 10-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Kenneth M.
REGISTRATION NUMBER: 34,174
REFERENCE/DOCKET NUMBER: 00081.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2719
TELEFAX: 510-655-3542
TELEX: N/A
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-207-901-50

Query Match 65.9%; Score 11.2; DB 1; Length 50;
Best Local Similarity 62.5%; Pred. No. 6.5e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgyttgcttcnacyt 16
||:| | | :| | | :| |
DB 21 TGTTAGCATCTCTT 36

RESULT 8
US-08-454-557C-74/C
Sequence 74, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-74

Query Match 65.9%; Score 11.2; DB 3; Length 53;
Best Local Similarity 62.5%; Pred. No. 6.5e+02;

Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 gytngcrtcnactg 17
1:1 11:11 11:11
Db 32 GCTATGATCACTTG 17

RESULT 9

US-08-340-426D-74/C

; Sequence 74, Application US/08340426D
; Patent No. 5948634

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/340,426D

FILING DATE: 14-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:

LENGTH: 53 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

US-08-340-426D-74

Query Match 65.9%; Score 11.2; DB 4; Length 53;
Best Local Similarity 62.5%; Pred. No. 6.5e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 gytngcrtcnactg 17
1:1 11:11 11:11
Db 32 GCTATGATCACTTG 17

RESULT 10

US-08-450-673C-74/C

; Sequence 74, Application US/08450673C
; Patent No. 5948888

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/450,673C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:

LENGTH: 53 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

US-08-450-673C-74

Query Match 65.9%; Score 11.2; DB 4; Length 53;
Best Local Similarity 62.5%; Pred. No. 6.5e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 gytngcrtcnactg 17
1:1 11:11 11:11
Db 32 GCTATGATCACTTG 17

RESULT 11

PCT-US95-17111A-74/C

; Sequence 74, Application PC/TUS9517111A

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

TITLE OF INVENTION: Neutral Thread Protein Gene Expression and

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/17111A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/340,426

FILING DATE: 14-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:

LENGTH: 53 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-74

Query Match 65.9%; Score 11.2; DB 6; Length 53;
Best Local Similarity 62.5%; Pred. No. 6.5e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 2 tgytngcrtcnacyt 17
1:1 11:11 11:11
Db 32 GCTATCATCATCACTG 17

RESULT 12
US-08-564-955-60
Sequence 60, Application US/08564955
Patent No. 5811238
GENERAL INFORMATION:
APPLICANT: STEMMER, WILLEM P.C.
APPLICANT: CRAMER, ANDREAS M.
TITLE OF INVENTION: METHODS FOR GENERATING POLYNUCLEOTIDES
TITLE OF INVENTION: HAVING DESIRED CHARACTERISTICS BY ITERATIVE SELECTION AND
TITLE OF INVENTION: RECOMBINATION
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: U.S.A.
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/564,955
FILING DATE: 30-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,874
FILING DATE: 30-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02126
FILING DATE: 17-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: DUNN, TRACY J.
REGISTRATION NUMBER: 34,587
REFERENCE/DOCKET NUMBER: 16528J-014611US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-564-955-60

Query Match 65.9%; Score 11.2; DB 2; Length 54;
Best Local Similarity 62.5%; Pred. No. 6.5e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 1 tgytngcrtcnacyt 16

Db 7 TCCTGCGCATCCACT 22
11:1 11:11 11:1

RESULT 13
US-08-564-955-61
Sequence 61, Application US/08564955
Patent No. 5811238
GENERAL INFORMATION:
APPLICANT: STEMMER, WILLEM P.C.
APPLICANT: CRAMER, ANDREAS M.
TITLE OF INVENTION: METHODS FOR GENERATING POLYNUCLEOTIDES
TITLE OF INVENTION: HAVING DESIRED CHARACTERISTICS BY ITERATIVE SELECTION AND
TITLE OF INVENTION: RECOMBINATION
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: U.S.A.
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/564,955
FILING DATE: 30-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,874
FILING DATE: 30-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02126
FILING DATE: 17-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: DUNN, TRACY J.
REGISTRATION NUMBER: 34,587
REFERENCE/DOCKET NUMBER: 16528J-014611US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-564-955-61

Query Match 65.9%; Score 11.2; DB 2; Length 57;
Best Local Similarity 62.5%; Pred. No. 6.6e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 1 tgytngcrtcnacyt 16
11:1 11:11 11:1
Db 7 TCCTGCGCATCCACT 22

RESULT 14
US-08-537-874-58
Sequence 58, Application US/08537874
Patent No. 5830721
GENERAL INFORMATION:
APPLICANT: STEMMER, WILLEM P.C.
APPLICANT: CRAMER, ANDREAS
TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation

```

; TITLE OF INVENTION: and Reassembly
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,874
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP PCT/US95/02126
; FILING DATE: 17-FEB-1995
; APPLICATION NUMBER: US 08/198,431
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018097-014610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-537-874-58

```

```

Query Match          65.9%; Score 11.2; DB 3; Length 57;
Best Local Similarity 62.5%; Pred. No. 6.6e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 tgyttingcrtcnacyt 16
    ||:| ||:| ||:|
Db 7 TGCTGCGCATCCACCT 22

```

```

RESULT 15
US-08-537-874-59
; Sequence 59, Application US/08537874
; Patent No. 5830721
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.
; TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,874
; FILING DATE:

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP PCT/US95/02126
; FILING DATE: 17-FEB-1995
; APPLICATION NUMBER: US 08/198,431
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018097-014610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-537-874-59

```

```

Query Match          65.9%; Score 11.2; DB 3; Length 57;
Best Local Similarity 62.5%; Pred. No. 6.6e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 tgyttingcrtcnacyt 16
    ||:| ||:| ||:|
Db 7 TGCTGCGCATCCACCT 22

```

```

Search completed: June 4, 2000, 16:09:39
Job time: 28058 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:26 ; Search time 1236.38 Seconds

(without alignments)
-15.736 Million cell updates/sec

Title: US-09-164-714-4

Perfect score: 20
Sequence: 1 gcngaygcncarcntaygt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_com:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pl4:*
11: gb_pl5:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl1:*
27: em_pl2:*
28: em_pl3:*
29: em_pl4:*
30: em_un:*
31: em_vl:*
32: gb_hg1:*
33: gb_hg2:*
34: gb_hg3:*
35: gb_hg4:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pl4:*
41: gb_hg3:*
42: gb_hg4:*
43: gb_hg5:*
44: gb_hg6:*

45: gb_hg7:*
46: em_hg1:*
47: em_hg2:*
48: em_hg3:*
49: em_hg4:*
50: gb_pl3:*
51: gb_pl4:*
52: gb_hg8:*
53: gb_hg9:*
54: gb_hg10:*
55: gb_hg11:*
56: gb_hg12:*
57: gb_hg13:*
58: gb_hg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	12.2	61.0	32	5	A69250	A69250 Sequence 8
2	12.2	61.0	32	5	A71927	A71927 Sequence 8
3	11.6	58.0	30	5	A17053	A17053 Oligonucleo
4	11.6	58.0	30	5	A17453	A17453 Oligonucleo
5	11.6	58.0	30	5	AR014428	AR014428 Sequence
6	11.6	58.0	30	5	111871	111871 Sequence 37
7	11.6	58.0	71	5	AR017688	AR017688 Sequence
8	11.4	57.0	21	5	AR030742	AR030742 Sequence
9	11.4	57.0	21	5	AR030743	AR030743 Sequence
10	11.4	57.0	39	5	AR022077	AR022077 Sequence
11	11.4	57.0	39	5	192867	192867 Sequence 41
12	11.4	57.0	60	5	A19037	A19037 Nucleotide
13	11.4	57.0	60	5	A19038	A19038 Nucleotide
14	11.4	57.0	75	5	125131	125131 Sequence 21
15	11.4	57.0	75	11	HS091211	HS091211 Homo sapien
16	11.4	57.0	90	5	125143	125143 Sequence 39
17	11.2	56.0	30	34	DRORRN	K01289 D.melanogas
18	11.2	56.0	52	14	SYNANVA4	M60112 Avian neovl
19	11.2	56.0	76	16	AB003918	AB003918 HepaticLis
20	11.2	56.0	85	5	A79759	A79759 Sequence 7
21	11.2	56.0	85	5	A79760	A79760 Sequence 8
22	11.2	56.0	90	12	RATVPN	K02434 Rat vasopre
23	11.2	56.0	90	12	RATVPNP	K01495 Rat prepro
24	11.2	56.0	97	5	114127	114127 Sequence 41
25	11.2	55.0	25	5	111740	111740 Sequence 11
26	11.2	55.0	27	5	AR039342	AR039342 Sequence
27	11.2	55.0	31	5	E02911	E02911 Oligonucleo
28	11.2	55.0	32	5	AR003523	AR003523 Sequence
29	11.2	55.0	32	5	162417	162417 Sequence 23
30	11.2	55.0	34	5	186938	186938 Sequence 27
31	11.2	55.0	35	5	119619	119619 Sequence 8
32	11.2	55.0	40	5	E06658	E06658 DNA probe t
33	11.2	55.0	43	5	100832	100832 Sequence 3
34	11.2	55.0	43	5	100841	100841 Sequence 4
35	11.2	55.0	43	5	106027	106027 Sequence 1
36	11.2	55.0	46	5	117466	117466 Sequence 5
37	11.2	55.0	49	12	MUSTGHP	M22386 Mouse 19 ge
38	11.2	55.0	53	5	AR062116	AR062116 Sequence
39	11.2	55.0	56	5	A48236	A48236 Sequence 15
40	11.2	55.0	72	14	AF167289	AF167289 Cloning v
41	11.2	55.0	79	5	AR042663	AR042663 Sequence
42	11.2	55.0	79	5	AR064796	AR064796 Sequence
43	11.2	55.0	84	11	HS243870	A0245870 Homo Sapi
44	11.2	55.0	94	9	H0MFAH11	L14667 Human tumor
45	10.8	54.0	57	5	A18403	A18403 signal sequ

ALIGNMENTS

OY 1 gcnaygncarcnclaygt 20
 |||||
 DB 39 GATGAGGCCGCGACGT 20

RESULT 15
 HSU91211 75 bp mRNA PRI 03-JUL-1997
 LOCUS HSU91211/c Homo sapiens T-cell receptor delta chain (TCRDV3J2) mRNA, partial
 DEFINITION cds.
 ACCESSION U91211
 VERSION U91211.1 GI:2239828
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 75)
 AUTHORS Holmeier,W., Witthoft,T., Hennemann,A., Winter,H.S. and
 Kagnoff,M.F.
 TITLE The TCR-delta repertoire in human intestine undergoes
 JOURNAL characteristic changes during fetal to adult development
 MEDLINE J. Immunol. 158 (12), 5632-5641 (1997)
 REFERENCE 97334214
 2 (bases 1 to 75)
 AUTHORS Holmeier,W., Witthoft,T., Hennemann,A., Harland,S.W. and
 Kagnoff,M.F.
 TITLE Direct Submission
 JOURNAL Submitted (27-FEB-1997) Department of Medicine, University of
 Frankfurt, Theodor-Stern Kai #7, Frankfurt 60590, Germany
 FEATURES
 Location/Qualifiers
 source
 1..75
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DOI363"
 /chromosome="14"
 /tissue_type="small intestine"
 /dev_stage="one day old"
 <1..>75
 /gene="TCRDV3J2"
 <1..>75
 /gene="TCRDV3J2"
 <1..>75
 /gene="TCRDV3J2"
 /note="rearranged; contains CDR3 domain, 13 amino acids
 (calculation according to: Rock,E.
 J. Exp. Med.179:323-328,1994)"
 /codon_start=1
 /product="T-cell receptor delta chain"
 /protein_id="AAC51428.1"
 /db_xref="GI:2239829"
 /translation="TEDSATYYCASALTGCGDALTRQLF"

BASE COUNT 15 a 20 c 21 g 19 t
 ORIGIN

Query Match 57.0%; Score 11.4; DB 11; Length 75;
 Best Local Similarity 60.0%; Pred. No. 3.4e+04;
 Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 OY 1 gcnaygncarcnclaygt 20
 |||||
 DB 35 GCCGACGCGACGAGT 16

Search completed: June 4, 2000, 16:05:28
 Job time: 2780 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:23:59 ; Search time 322.35 Seconds
(without alignments)
15.523 Million cell updates/sec

Title: US-09-164-714-4

Sequence: 1 gongaygencarcntaygt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12.2	61.0	32	1	V49438 Primer AB028 for B
2	12	60.0	50	1	c-fos position 272
3	11.6	58.0	40	1	Oligonucleotide ge
4	11.6	58.0	40	1	Oligonucleotide ge
5	11.6	58.0	53	1	C. limi endoglucan
6	11.6	58.0	71	1	L-selectin family
7	11.4	57.0	21	1	Human cathepsin K
8	11.4	57.0	21	1	Human cathepsin K
9	11.4	57.0	39	1	5'-Primer for BR96
10	11.4	57.0	47	1	PCR primer C3634CC
11	11.4	57.0	63	1	Drosophila melanog
12	11.4	57.0	63	1	Codon optimised s1
13	11.4	57.0	75	1	Codon optimised D.
14	11.4	57.0	90	1	Fragment B esteras
15	11.4	57.0	90	1	AAT optimised seq
16	11.4	57.0	90	1	Oligomer for a cod
17	11.2	56.0	97	1	Sequence of synthe
18	11	55.0	25	1	CFTR panhandle PCR
19	11	55.0	27	1	Human IL-2 recepto
20	11	55.0	29	1	PCR primer p1 to a
21	11	55.0	32	1	Human cDNA clone #
22	11	55.0	32	1	Lambda 37225-3725
23	11	55.0	35	1	Probe for human ba
24	11	55.0	39	1	PCR primer for PET
25	11	55.0	39	1	PCR primer Sq2258
26	11	55.0	40	1	Mycoplasma plasm/V
27	11	55.0	40	1	PCR primer for hum
28	11	55.0	43	1	Probe for 2,5-dike
29	11	55.0	43	1	Oligonucleotide pr
30	11	55.0	43	1	Murine osteoproteg
31	11	55.0	44	1	p1 VH cDNA primer
32	11	55.0	46	1	Murine osteoproteg
33	11	55.0	49	1	Hepatoma AS-30D Ty

35	11	55.0	53	1	063430	Lambda phage repl1
36	11	55.0	53	1	V70296	Murine osteoproteg
37	11	55.0	56	1	T16105	Plasmid pBR322 PCR
38	11	55.0	75	1	Q11149	Probe GTR-1 based
39	11	55.0	81	1	X32352	Artificial NS4 mos
40	11	55.0	83	1	063427	Lambda phage repl1
41	11	55.0	83	1	X32356	Artificial NS4 mos
42	10.6	53.0	19	1	V72009	Electronic perturb
43	10.6	53.0	21	1	Q35337	PCR primer HIV3B9.
44	10.6	53.0	23	1	V64049	Human leiomyoma DN
45	10.6	53.0	24	1	Q75804	Trimer oligonucleo

ALIGNMENTS

RESULT 1	
V49438	
ID	V49438 standard; DNA; 32 BP.
AC	V49438; 1998 (first entry)
DT	28-OCT-1998
DE	Primer AB028 for BRSV protein G gene.
KW	Multivalent vaccine; pathogen; respiratory disease; digestive disease;
KW	cattle; bovine herpesvirus; BHV; bovine respiratory syncytial virus;
KW	BRSV; bovine viral diarrhoea virus; BVDV; bovine parainfluenza virus;
KW	BPTV; vector; glycoprotein; primer; PCR; amplification; ss.
OS	Synthetic.
OS	Bovine respiratory syncytial virus.
PN	FR2751229-A1.
PD	23-JAN-1998.
PF	19-JUL-1996; 009403.
PR	19-JUL-1996; FR-009403.
PA	(INMR) RHONE MERIEUX SA.
PI	Audomnet JCF, Baudu P, Bouchardon A, Riviere MEA;
DR	WPI: 98-112829/11.
PT	Multi-valent polynucleotide vaccines against bovine pathogens
PT	consist of at least 3 plasmids able to express protective antigens
PT	from specified viruses
PS	Example 10; Page 13; 33p; French.
CC	The invention relates to a multivalent vaccine for protecting cattle
CC	against several pathogens, especially pathogens associated with
CC	respiratory and digestive diseases. The pathogens are especially
CC	selected from bovine herpesvirus (BHV), bovine respiratory syncytial
CC	virus (BRSV), bovine viral diarrhoea virus (BVDV) or bovine parainfluenza
CC	virus (BPTV). The vaccines are preferably composed of polynucleotide
CC	sequences encoding 3 antigens, all as part of vectors.
CC	Primers V49438-V49439 were used to PCR amplify the BRSV strain 391-2
CC	protein G gene. The sequence was subcloned into the plasmid pVR1012
CC	to generate plasmid pAB012 for use in the vaccine.
SQ	Sequence 32 BP; 12 A; 11 C; 3 G; 6 T;
Query Match	61.0%; Score 12.2; DB 1; Length 32;
Best Local Similarity	61.1%; Pred. No. 3.6e+02;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;	
Db	1 gongaygencarcntaygt 18
	: : :
	7 GCAGATGTCACACCTAC 24
RESULT 2	
Q50214	
ID	Q50214 standard; DNA; 50 BP.
AC	Q50214;
DT	06-MAY-1994 (first entry)
DE	c-fos position 2721-2770 INS mutagenic oligonucleotide.
KW	HIV-1; p17gag; inhibition; gag; M1; M2; M3; M4; silent mutation; mRNA;
KW	vector; p17; point mutation; p17M1234; Hlat cells; gene replacement;
KW	inhibitory/instability signal; INS; stability; utilisation; vaccine;
KW	interferon; interleukin; fos proto-oncogene protein; growth factor;
KW	env; attenuated; AIDS; AIDS-related disease; latent infection;
KW	gene therapy; Human immunodeficiency virus type 1; p24;

KM long terminal repeat; LTR; ss.
 OS Synthetic.
 PN MO3320212-A.
 PD 14-OCT-1993.
 PE 29-MAR-1993.
 PR 27-MAR-1992; US-858747.
 PA (USSH) US SEC DEPT HEALTH.
 PI Felber BK, Pavlakis GN;
 DR WPI: 93-336919/42.
 PT Eliminating inhibitory-instability regions in mRNA to improve
 PT stability and expression - by making multiple point mutations
 PT within A-T rich regions
 PS Example 3; Page 70; 117pp; English.
 CC The sequences given in Q50202-70 are oligonucleotides which were
 CC used to illustrate the method of the invention for the mutagenesis of
 CC inhibitory/instability signals (INS). Mutation in an INS produced
 CC using oligonucleotides such as these, increases the stability and/or
 CC utilisation of mRNA without changing its protein coding capacity, or
 CC if the sequence is changed, its function is maintained. Other genes
 CC encoding such mRNA molecules include growth factor, interferon,
 CC interleukin, fos proto-oncogene protein and HIV env and gag gene
 CC products. Nucleic acid constructs in which INS function has been
 CC impeded, can be used as vaccines, esp. against AIDS and AIDS-related
 CC diseases by preventing HIV from establishing a latent infection, as
 CC is possible using the INS, and thus escaping immune system
 CC surveillance. The constructs may also be used in gene therapy for
 CC gene replacement by homologous recombination with a target gene in
 CC situ. See also q50200-02.
 SO Sequence 50 BP; 16 A; 13 C; 13 G; 8 T;

	Query Match	60.0%;	Score 12;	DB 1;	Length 50;
	Best Local Similarity	63.2%;	Pred. No. 4.9e+02;		
	Matches 12;	Conservative	2;	Mismatches 5;	Indels 0;
0y	2 cngaygcncarcnclatg	20			
	1 1 1 1 1 1 1 1 1 1				
db	17 CTGAAGGACAGCCCATACGT	35			

RESULT 3
 ID Q38054/C
 AC Q38054;
 DT 07-JUL-1993 (first entry)
 DE Oligonucleotide gell1, for prodn. of synthetic gelonin gene.
 KW Seed, toxin; plant; cloning; ribosomal; protein synthesis; ss;
 KW Gelonium multiflorum.
 OS Synthetic.
 PN W09305168-A.
 PD 18-MAR-1993.
 PE 21-AUG-1992; US-755949.
 PR 06-SEP-1991; US-755949.
 PA (RERE-) RES DEV FOUND.
 PI Beattie KL, Rosenblum MG;
 WP1: 93-100990/12.
 PT Synthetic DNA encoding gelonin plant toxin - provides nucleotide
 PT sequence for synthetic gene for prodn. and cloning
 PS Example 4; Fig 5; 45pp; English.
 CC The synthetic gelonin gene based on the sequence of Gelonium
 CC multiflorum gelonin gene prepd. by synthesizing a number of
 CC oligonucleotides corresp. to fragments of the gelonin gene and
 CC annealing and ligating to assemble the intact gene. The
 CC oligonucleotides were designed to contain a codon triplet for each
 CC amino acid in the corresp. gelonin fragment. Gelonin is a ribosomal-
 CC inactivating plant toxin which inhibits protein synthesis. The
 CC synthetic form of gelonin provides a plentiful, reproducible source
 CC of gelonin which may be modified.
 CC See also Q38041-82.
 Q3 Sequence 40 BP; 15 A; 12 C; 9 G; 4 T;

Query Match 58.0%; Score 11.6; DB 1; Length 40;

```

Best Local Similarity 57.9%; Pred. No. 7.9e+02;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 gcngaygcncarcntayg 19
      ||||| ||| | :|
Db      19 GCCGATGCCACAGCGGTTCG 1

```

```

RESULT      4
ID          038072 standard; DNA; 40 BP.
AC          038072;
DT          07-JUL-1993 (first entry)
DE          Oligonucleotide gel29, for prodn. of synthetic gelonin gene.
KW          Seed; toxin; plant; cloning; ribosomal; protein synthesis; ss;
KW          Gelonium multiflorum.
OS          Synthetic.
PN          WO9305168-A.
PD          18-MAR-1993.
PF          21-AUG-1992; U07066.
PR          06-SEP-1991; US-755949.
PA          (RERE-) RES DEV FOUND.
PI          Beattie RL, Rosenblum MG;
DR          WPI: 93-100990/12.
PT          Synthetic DNA encoding gelonin plant toxin - provides nucleotide
PT          sequence for synthesizing gene for prodn. and cloning
PS          Example 4; Fig 5; 45pp; English.
CC          The synthetic gelonin gene based on the sequence of Gelonium
CC          multiflorum gelonin was prepd. by synthesizing a number of
CC          oligonucleotides corresp. to fragments of the gelonin gene and
CC          annealing and ligating to assemble the intact gene. The
CC          oligonucleotides were designed to contain a codon triplet for each
CC          amino acid in the corresp. gelonin fragment. Gelonin is a ribosomal-
CC          inactivating plant toxin which inhibits protein synthesis. The
CC          synthetic form of gelonin provides a plentiful, reproducible source
CC          of gelonin which may be modified.
SQ          See also 038041-82.
SQ          Sequence 40 BP; 7 A; 11 C; 12 G; 10 T;

```

	Query Match	58.0%;	Score 11.6;	DB 1;	Length 40;
	Best Local Similarity	57.9%;	Pred. No. 7.9e025;		
	Matches 11;	Conservative	3;	Mismatches 5;	Indels 0;
Oy	1 gcngaygcncarcrcntayg	19			
	: :	1:			
Db	2 GCCGATGCGCACCGCGTCG	20			

RESULT 5
ID V29675 standard; DNA; 53 BP.
AC V29675;
DT 01-SEP-1998 (first entry)
DE C. fimi endoglucanase A (Cena) gene amplifying primer 15, 3'SAENH,
KW starch; liquefaction; sweetener; enzymes hybrid; endoglucanase; enzyme
KW cellulose binding domain; CBD; starch processing; alpha-amylase;
KW saccharilic acid; Cena; PCR primer; ss.
OS Synthetic.
OS Cellulomonas fimi.
PN MO9816633-A1.
PD 23-APR-1998.
PE 13-OCT-1997; DK0448.
PR 11-OCT-1996; DK-001130.
PR (NOVO) NOVO-NORDISK AS.
PI Bisgaardifantzen H, Bjornvad M, Pedersen S, Schulein M;
PR WPI; 98-251283/22.
PR Liquefaction of starch for, e.g. production of sweeteners -
PR comprises use of enzyme hybrids including cellulose binding domain
PR for starch
PS Example 7; Page 39; 83pp; English.
CC This primer is used for the PCR amplification of the Cellulomonas fimi
CC endoglucanase A (Cena) gene fragment encoding the cellulose binding


```

CC domain (CBD). This is used in the construction of enzyme hybrids for
CC liquefaction of starch. The enzyme hybrids contain amino acid sequences
CC of alpha-amylase linked to a CBD. The starch is liquefied by treating,
CC in aqueous medium, with such an enzyme hybrid. A recombinant expression
CC vector comprising a construct containing isolated DNA encoding enzyme
CC hybrids with amylolytic activity, promoter and stop signals can be used
CC to transform host cells for the production of the recombinant enzyme
CC hybrids. The enzyme hybrids are useful in industrial starch processing
CC especially for the production of sweeteners. Hybrid enzymes have altered
CC affinity for substrate and increased activity, resulting in at least 1 of
CC reduced calcium ion dependence, reduced formation of Mallard reaction
CC products and reduced effect of alpha-amylase on subsequent
CC saccharification.
SQ Sequence 53 BP; 11 A; 13 C; 16 G; 13 T;

Query Match 58.0%; Score 11.6; DB 1; Length 53;
Best Local Similarity 57.9%; Pred. No. 8.2e+02;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 gengaygcncarcntaygt 19
   |||||  |||||  |||
Db 51 GCCGACGACACGCCACG 33

RESULT 6
TS7762/c
ID TS7762 standard; RNA; 71 BP.
AC TS7762;
DE 21-NOV-1997 (first entry)
KW L-selectin family III SELEX ligand F13.57.
KW Identification; ligand; lectin; SELEX; wheat germ agglutinin; template;
KW Systematic Evolution of Ligands by Exponential enrichment; amplification;
KW primer; PCR; polymerase chain reaction; peritoneal inflammation; ss;
KW diabetes; lymphocyte trafficking disorder; glomerulonephritis; arthritis.
FS Synthetic.
FH Key location/Qualifiers
FT modified_base 1..71
FT /*tag= a
FT /mod_base= all C bases are 2' NH2-cytosine
FT /mod_base= all U bases are 2' NH2-uracil
PN MO9640703-A1.
PD 19-DEC-1996.
PE 05-JUN-1996; U09455.
PR 07-JUN-1995; US-479724.
PR 07-JUN-1995; US-472255.
PR 07-JUN-1995; US-472256.
PR 07-JUN-1995; US-477829.
PA (NEXS-) NEXSTAR PHARM INC.
PI Bridonnan P, Gold L, Hicke B, Parma DH;
DR WPI; 97-077252/07.
PT Identifying nucleic acid ligands that bind lectin(s) esp.
PT selectin(s) - by partitioning the ligands from a mixture of nucleic
PT acids
PS Claim 35; Page 148; 255pp; English.
CC The invention relates to the identification of nucleic acid ligands to
CC a lectin using the Systematic Evolution of Ligands by Exponential
CC enrichment (SELEX) method. The sequences T57740-T57790 represent RNA
CC ligands isolated by the method which bind to L-selectin. The L-selectin
CC ligands were isolated from a DNA template containing 40 random
CC nucleotides flanked by fixed 5' and 3' sequences (T58043), which was
CC amplified using the primers T58044-5. The ligands fall into 13 families
CC along with a group of unrelated 'orphan' ligands. This ligand binds
CC L-selectin with a Kd of 75 nM at room temperature. The ligands are
CC especially useful in the treatment of peritoneal inflammation, diabetes,
CC lymphocyte trafficking disorders, glomerulonephritis, arthritis, etc.
SQ Sequence 71 BP; 19 A; 15 C; 26 G; 11 U;

Query Match 58.0%; Score 11.6; DB 1; Length 71;
Best Local Similarity 57.9%; Pred. No. 8.4e+02;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 2 cngaygcncarcntaygt 20
   |||||  |||||  |||||
Db 36 CAGACCGCTACTTATGT 18

RESULT 7
V09686/c
ID V09686 standard; DNA; 21 BP.
AC V09686;
DE 20-JUL-1998 (first entry)
DE Human cathepsin K gene exon 5 reverse PCR primer 4R.
KW Cathepsin K; human; osteoporosis; periodontal disease;
KW Paget's disease; Gaucher's disease; Alzheimer's disease;
KW central nervous system inflammation; hyperparathyroidism;
KW bone degradation; dental implant degradation; metastasis; tumour;
KW diagnosis; therapy; marker; PCR; primer; ss.
FS Synthetic.
FH Homo sapiens.
FH EP-812916-A2.
PD 17-DEC-1997.
PE 19-MAY-1997; 303395.
PE 26-AUG-1996; US-026083.
PR 14-JUN-1996; US-019942.
PR 17-JUN-1996; US-020273.
PR 26-AUG-1996; WO-014026.
PA (HUMA-) HUMAN GENOME SCI INC.
PA (GENO-) INST GENOMIC RES.
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PI Adams MD, Blake JA, Debouck CM, Drake FH, Fitzgerald LM,
PI Fraser CM, Gowen M, Hastings GA, Kirkness EF, Lee NH,
PI Rodd J;
DR WPI; 98-034977/04.
PT DNA encoding human cathepsin K - useful for diagnosing and treating
PT diseases associated with cathepsin K e.g. osteoporosis, bone
PT degradation, metastatic tumours, etc
PS Example 1; Page 50; 84pp; English.
PS Oligonucleotide 4R comprises a reverse PCR primer from exon 5
CC (see V09670) of the human cathepsin K gene (see V09660). PCR
CC primers (see V09679-90) to adjacent exons of the cathepsin K
CC gene were used in the amplification of human genomic DNA. DNA
CC sequencing of intron-exon boundaries allowed sequencing of the
CC cathepsin genomic DNA. DNA encoding human cathepsin K is useful
CC for the diagnosis and treatment of e.g. osteoporosis, periodontal
CC disease, Paget's disease, Gaucher's disease, CNS inflammation,
CC Alzheimer's disease, hyperparathyroidism, bone degradation,
CC metastatic tumours, and degradation of bone implants and
CC prostheses, especially dental implants.
SQ Sequence 21 BP; 5 A; 5 C; 5 G; 6 T;

Query Match 57.0%; Score 11.4; DB 1; Length 21;
Best Local Similarity 60.0%; Pred. No. 9.5e+02;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 gengaygcncarcntaygt 20
   |||||  |||||  |||||
Db 21 GAAGATGCTACCATATGT 2

RESULT 8
V09687
ID V09687 standard; DNA; 21 BP.
AC V09687;
DE 20-JUL-1998 (first entry)
DE Human cathepsin K gene exon 5, 6 forward PCR primer 5+6F.
KW Cathepsin K; human; osteoporosis; periodontal disease;
KW Paget's disease; Gaucher's disease; Alzheimer's disease;
KW central nervous system inflammation; hyperparathyroidism;
KW bone degradation; dental implant degradation; metastasis;
KW diagnosis; therapy; marker; PCR; primer; ss.
FS Synthetic.
FH Homo sapiens.
FH EP-812916-A2.

```


DR MP1: 94.42108/230.
PT Heterologous signal sequences for secretion of insect controlling
PS proteins - useful to protect plants from insect pests
PS Example 2: Page 14, 63pp, English.
CC The sequences given in Q67700-10 are fragments which were used in the
CC construction of eight heterologous signal sequence-codon optimised AaIT
CC toxin gene cassettes. The cassettes were constructed in two pieces, a
CC "A" fragment unique for each construct, consisting of DNA coding for one
CC of eight heterologous signal sequences, plus the amino terminal portion
CC of the toxin coding region, and an "A" fragment which is the same for
CC each construct and encodes the remainder of the toxin coding gene. Each
CC of the fragments "A" and "B" were synthesised by annealing a pair of
CC oligomers containing a 15 bp overlap to complete the double stranded
CC molecule. Fusion constructs such as these, may be introduced into an
CC insect virus, such as the baculovirus AcMNPV. The insertion of the AaIT
CC gene and the heterologous signal sequence into a baculovirus results in
CC the expression and secretion of the toxin. A susceptible insect which
CC ingests such a modified baculovirus will cease feeding on plants due to
CC toxin-induced paralysis at an earlier time than an insect which ingests
CC a wild-type baculovirus, thus reducing crop damage.
SQ Sequence 90 BP; 17 A; 25 C; 28 G; 20 T;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:52:59 ; Search time 4521.53 Seconds
(without alignments)
17.929 Million cell updates/sec

Title: US-09-164-714-4

Perfect score: 20
Sequence: 1 gengaygcncarcntaygt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 08
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gd_est1:*
21: gd_est2:*
22: gd_est3:*
23: gd_est4:*
24: gd_est5:*
25: gd_est6:*
26: gd_est7:*
27: gd_est8:*
28: gd_est9:*
29: gd_est10:*
30: gd_est11:*
31: gd_est12:*
32: gd_est13:*
33: gd_est14:*
34: gd_est15:*
35: gd_est16:*
36: gd_est17:*
37: gd_est18:*
38: gd_est19:*
39: gd_est20:*
40: gd_est21:*
41: gd_est22:*
42: gd_est23:*
43: gd_est24:*
44: gd_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result

Query

WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:584819
 Putative full length read
 vector to vector length is 168
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 67.
 Location/Qualifiers

FEATURES

source

1.75
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:1053243"
 /dev_stage="8 weeks"
 /dev_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site.1: EcoRI; Site.2: NotI; Tissue obtained
 from 8 week old mouse. Colon was harvested 72 hours after
 irradiation with 1400 Gys. 1st strand cDNA was primed
 with a Not I - oligo(dT) primer
 [5'TGTTCAGATCTGAGTGGAGCGCGCTTTTCTTTTCTTTTCTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI
 adaptors [AATTCGATCCTTG], digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT73
 vector. Library constructed by Bob Barstead."

BASE COUNT
 21 a 19 c 20 g 15 t

Query Match 63.0% Score 12.6; DB 36; Length 75;
 Best Local Similarity 70.6%; Pred. No. 2.4e+03;

Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 4 gayencarcntaygt 20
 |||||
 Db 18 GAGGCCCAACCTATGCT 34

RESULT 3

D63753/c

LOCUS D63753 98 bp mRNA EST 18-NOV-1997
 DEFINITION D63753 Human chromosome 11q23 mRNA (M.Katoh) Homo sapiens CDNA

ACCESSION D63753
 VERSION D63753.1 GI:2326278

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

AUTHORS

1 (bases 1 to 98)
 Katoh, M., Nakagawa, Y., Yawata, T., Kumano, S., Kobayashi, E.,
 Kurimasa, A., Kugoh, H. and Oshimura, M.

TITLE

Cosmids and transcribed sequences from chromosome 11q23

JOURNAL

Jpn. J. Hum. Genet. 40 (4), 307-317 (1995)

MEDLINE

97004453

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1394933.
 Contact: Motoonobu Katoh
 Faculty of Medicine, Department of Molecular and Cell Genetics
 Tottori University
 86 Nishimachi, Yonago, Tottori 683, Japan
 Email: mkatoh@grape.med.tottori-u.ac.jp.
 Location/Qualifiers

1.98
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

/map="11q23"
 /clone="pEX4A-3"
 /clone_lib="Human chromosome 11q23 mRNA (M.Katoh)"
 35 a 21 c 21 g 21 t

Query Match 63.0% Score 12.6; DB 35; Length 98;
 Best Local Similarity 70.6%; Pred. No. 2.7e+03;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 4 gayencarcntaygt 20
 |||||
 Db 38 GAGGCTCATCCATATGCT 22

RESULT 4

AA177604

LOCUS AA177604 70 bp mRNA EST 16-FEB-1997
 DEFINITION MC25910.r1 Soares mouse 3BDMS Mus musculus cDNA clone IMAGE:622146

5' similar to SW:NU4M.MOUSE P03911 NADH-UBIQUINONE OXIDOREDUCTASE
 CHAIN 4; mRNA sequence.

ACCESSION

AA177604

VERSION

AA177604.1 GI:1758838

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

AUTHORS

1 (bases 1 to 70)
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, V., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE

The WashU-HMI Mouse EST Project

JOURNAL

Unpublished (1996)

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1393548.
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:382970

Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES

source

1.70
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:622146"
 /clone_lib="Soares mouse 3BDMS"
 /sex="male"
 /tissue_type="Spleen"
 /dev_stage="4 weeks"
 /dev_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTTCAGATCTGAGTGGAGCGCGCTGTTTCTTTTCTTTTCTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through
 three rounds of normalization, and was constructed by
 Bento Soares and M.Fatima Bonaldo."

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 88)
COMMENT NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Apr 14, 1993 this sequence version replaced g1:693382.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,
M.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www.bio.llnl.gov/bdrip/image/image.html

FEATURES
Source
Insert Length: 1746 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 88
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:916733"
/clone_lib="NCI-CGAP_Pt2"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: PAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected preneoplastic cells
histologically-determined to be prostatic intraepithelial
neoplasia 2 (PIN2) cells. Double-stranded cDNA was
ligated to EcoRI adaptors, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into PAMP10 by the UDG-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."

BASE COUNT
ORIGIN
21 a 30 c 23 g 14 t

Query Match 60.0%; Score 12; DB 35; Length 88;
Best Local Similarity 63.2%; Pred. No. 5.6e+03;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 gcgaagcncarcantaygt 19
|||:|||||:|||||:|||||:
Db 17 GCTGATGCACACGCTACG 35

RESULT 8
AI094466/c 67 bp mRNA EST 10-NOV-1998
LOCUS ou87f09.s1 Soares_NSF_F8_9W_OT_PA_P-S1 Homo sapiens CDNA clone
DEFINITION IMAGE:1634825 3' similar to contains MER22.t2 MER22 repetitive
element; , mRNA sequence.
ACCESSION AI094466 GI:3433442
VERSION AT094466.1 GI:3433442
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 67)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL On Jan 19, 1998 this sequence version replaced g1:2284745.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LINL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 431 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 53.
Location/Qualifiers
1. 67
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1634825"
/clone_lib="Soares_NSF_F8_9W_OT_PA_P-S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148672-149255, 15002 -
150407, 151176-152327 Soares NB2HP8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-726407, 739080-740999 subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT
ORIGIN
23 a 13 c 18 g 13 t

Query Match 58.0%; Score 11.6; DB 42; Length 67;
Best Local Similarity 57.9%; Pred. No. 8.6e+03;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 cngaagcncarcantaygt 20
|||:|||||:|||||:|||||:
Db 59 CCGTGCACACGCTACGCT 41

RESULT 9
AA962216/c 76 bp mRNA EST 15-MAY-1998
LOCUS ou80603.s1 NCI CGAP Kid3 Homo sapiens CDNA clone IMAGE:1553524 3'
DEFINITION similar to TR:P97876 P97876 JUN DIMERIZATION PROTEIN 1 JDP-1. ;,
mRNA sequence.
ACCESSION AA962216
VERSION AA962216 GI:3134380
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 76)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL On Sep 12, 1996 this sequence version replaced g1:1402297.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Seq primer: -40m13 fwd, ET from Amersham
 High quality sequence stop: 1.

FEATURES

source

```
1. 76
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:153524"
/clone_lib="NCI-CGAP_Kid3"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. mRNA
source: 2 pooled kidneys. Library went through one round
of normalization. Library constructed by Bento Soares and
M. Fatima Bonaudo."
BASE COUNT      15 a      23 c      22 g      16 t
ORIGIN
```

Query Match

Best Local Similarity 58.0%; Score 11.6; DB 40; Length 76;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 gccgagcncarcctatg 19
 11 | | | | | 11 | | |
 Db 44 GCTGCTCTGTCCTATG 26

RESULT 10

LOCUS

AI965940 85 bp mRNA EST 06-DEC-1999
 sc24b04.y1 Gm-c1013 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-c1013-1544 5' similar to FR:004918 004918 CHLOROPHYLL A-B

BINDING PROTEIN OF LHCI TYPE III PRECURSOR ; mRNA sequence.
 AI965940
 EST.

ACCESSION AI965940.1 GI:5760577
 VERSION
 KEYWORDS

SOURCE

ORGANISM soybean.
 glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
 Glycine.

REFERENCE

AUTHORS

1 (bases 1 to 85)
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
 Bowers, J., Person, B., Swaller, T., Gildons, M., Pape, D., Harvey, N.,
 Schurr, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)

TITLE

JOURNAL

COMMENT

On May 18, 1998 this sequence version replaced gi:3138522.
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available through: Genome Systems, Inc. 4633 World

Parkway Circle St. Louis, Missouri 63134 For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or
info.genomesystems.com web site: www.genomesystems.com
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40RP from gibco
 High quality sequence stop: 1.

FEATURES

source

```
1. 85
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1013-1544"
/clone_lib="Gm-c1013"
/tissue_type="Whole seedlings, 2-3 week old seedlings,
greenhouse grown"
/lab_host="XL10-Gold"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from whole seedlings of 2-3 week old greenhouse grown
plants. The cDNA library was prepared using the Stratagene
pBluescript II XR cDNA library construction kit.
Complementary DNA was synthesized from mRNA using a primer
consisting of a poly (dT) sequence with a XhoI restriction
site. EcoRI adaptors were ligated to the blunt-ended cDNA
fragments followed by XhoI digestion. The cDNA fragments
were directionally cloned into the EcoRI-XhoI restriction
site of the pBluescript vector. The ligated cDNA fragments
were transformed into XL10-Gold host cells. This library
was constructed by Dr. Randy Shoemaker and Dr. John
Erpelting."
BASE COUNT      20 a      21 c      26 g      18 t
ORIGIN
```

Query Match

Best Local Similarity 58.0%; Score 11.6; DB 63; Length 85;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 gccgagcncarcctatg 19
 11 | | | | | 11 | | |
 Db 20 GCTGACCGCAGCATATG 38

RESULT 11

LOCUS

AA862553 100 bp mRNA EST 24-AUG-1998
 oh44b08.s1 NCI-CGAP_GC4 Homo sapiens cDNA clone IMAGE:1469463 3'
 similar to SW:DDX4_RAT 064060 DEAD BOX PROTEIN 4 ; mRNA sequence.
 AA862553
 AA862553.1 GI:2955032

ACCESSION
 VERSION
 KEYWORDS

SOURCE

ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homini; Hominidae; Homo.
 1 (bases 1 to 100)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1402122.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 1202 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amerisham
High quality sequence stop: 1.

FEATURES

source

```
1. 100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI_CGAP_GC4"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dt) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
```

BASE COUNT

30 a 18 c 19 g 33 t

ORIGIN

Query Match 58.0%; Score 11.6; DB 39; Length 100;
Best Local Similarity 57.9%; Pred. No. 9.7e+03;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 2 cngaygcncarcntaygt 20

Db 60 CAGATGCTCAACATGATGT 42

RESULT 12

T61825 36 bp mRNA EST 14-FEB-1995
LOCUS T61825/c
DEFINITION yb92804.s1 Stragene liver (#937224) Homo sapiens cDNA clone
IMAGE:78606 3' similar to gb:L11566 60S RIBOSOMAL PROTEIN L18
(HUMAN); mRNA sequence.

ACCESSION

T61825
T61825.1 GI:665068

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 36)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,
Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, N.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Insert size: 693
High quality sequence starts: 1 High quality sequence stops: 1
Source: IMAGE Consortium, LMB. This clone is available royalty-free
through LMB; contact the IMAGE Consortium (info@image.lmb.gov)
for further information. Trace considered overall poor quality
Insert length: 693 Std Error: 0.00
Seq primer: -21ml3
High quality sequence stop: 1.
Location/Qualifiers

source

```
1. 36
/organism="Homo sapiens"
/db_xref="GDB:498351"
/db_xref="taxon:9606"
/clone_lib="IMAGE:78606"
/clone_id="Stragene liver (#937224)"
/sex="male"
/age_stage="49 years old"
/lab_host="SOLR cells (kanamycin resistant)"
/notes="Organ: liver; Vector: pBluescript SK; Site_1:  
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. Hepatectomy from normal male caucasian. Average  
insert size: 1.1 kb; Uni-ZAP XR Vector; ~5' adaptor  
sequence: 5' GAATTCGCGACGAC 3' ~3' adaptor sequence: 5'  
CTGCGATTTTCTTTTCTTTTCTTTT 3"
```

BASE COUNT

3 a 9 c 15 g 7 t 2 others

ORIGIN

Query Match 56.0%; Score 11.2; DB 21; Length 36;
Best Local Similarity 62.5%; Pred. No. 1.2e+04;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 5 aycgncarcntaygt 20

Db 31 ACGCCAACCCTACCT 16

RESULT 13

A0025502 73 bp DNA GSS 14-OCT-1998
LOCUS A0025502/c
DEFINITION EP(X)1563-5prime Drosophila melanogaster EP line Drosophila
melanogaster genomic Sequence recovered from 5' end of P element,
genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 73)
Rehm, E.J. and Rubin, G.M.
The BDP gene disruption project: single EP element insertions
Unpublished (1998)
Contact: Gerald Rubin
Berkeley Drosophila Genome Project
University of California, Berkeley
LSA Building, Berkeley, CA 94720-3200, USA
Fax: 5106439947
Email: gerry@fruitfly.berkeley.edu

Sequence recovery method was inverse PCR.

Sequence orientation is forward strand relative to 5' end of P
element

The P element insertion position is base 66 in the 73 bases. This
insertion position refers to the first base of the 8 base target
recognition sequence.

Class: transposon-tagged.

FEATURES

source

```
1. 73
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_id="Drosophila melanogaster EP line"
/notes="Inverse PCR was performed on Drosophila  
melanogaster strains each of which contains a single EP  
transposable element insertion. (The generation of these  
insertion strains is described in North P, Szabo K, Bailey  
A, Laverly T, Rehm J, Rubin GM, Weigmann K, Milian M, Benes
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:39 ; Search time 244.64 Seconds
(without alignments)
10.627 Million cell updates/sec

Title: US-09-164-714-4

Perfect score: 20

Sequence: 1 gcngaycncarcrcntayt 20

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 segs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/6.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/PCTUS.COMB.seq:*
7: /cgn2_6/ptodata/2/lna/Backfills1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	60.0	50	4	US-08-850-049-18
2	12	60.0	50	4	US-08-050-478-18
3	11.8	59.0	68	5	US-08-469-318-99
4	11.8	59.0	68	5	US-08-469-318-100
5	11.8	59.0	68	6	PCT-US95-01185-99
6	11.8	59.0	68	6	PCT-US95-01185-100
7	11.6	58.0	30	1	US-07-734-225A-37
8	11.6	58.0	30	1	US-07-692-995B-37
9	11.6	58.0	30	2	US-08-488-457-37
10	11.6	58.0	33	2	US-08-366-953A-18
11	11.6	58.0	71	2	US-08-472-255A-89
12	11.6	58.0	71	2	US-08-479-724A-89
13	11.6	58.0	71	5	US-08-472-256B-89
14	11.6	58.0	71	6	PCT-US96-09455A-89
15	11.4	57.0	21	3	US-08-852-807-28
16	11.4	57.0	21	3	US-08-852-807-29
17	11.4	57.0	35	4	US-08-370-156-23
18	11.4	57.0	35	2	US-08-285-936-41
19	11.4	57.0	39	2	US-08-487-860-41
20	11.4	57.0	47	6	PCT-US93-1163B-11
21	11.4	57.0	63	6	PCT-US94-06079-41
22	11.4	57.0	75	1	US-08-009-265-21
23	11.4	57.0	86	4	US-08-477-527A-94
24	11.4	57.0	86	4	US-08-481-710-94
25	11.4	57.0	86	6	PCT-US96-09537-94
26	11.4	57.0	90	1	US-08-009-265-39
27	11.2	56.0	36	6	PCT-US95-15800-13

c 28	11.2	56.0	36	6	PCT-US95-15800-14	Sequence 14, Appl
c 29	11.2	56.0	97	1	US-07-750-080A-41	Sequence 41, Appl
30	11	55.0	25	1	US-07-786-902-11	Sequence 11, Appl
31	11	55.0	27	2	US-08-758-306-190	Sequence 190, Appl
32	11	55.0	32	1	US-08-330-535A-23	Sequence 23, Appl
33	11	55.0	32	2	US-08-688-145-20	Sequence 20, Appl
34	11	55.0	32	3	US-08-838-844-23	Sequence 23, Appl
35	11	55.0	34	2	US-08-170-290A-27	Sequence 27, Appl
36	11	55.0	35	1	US-08-125-618-8	Sequence 8, Appl
37	11	55.0	35	4	US-08-397-504-8	Sequence 8, Appl
c 38	11	55.0	39	3	US-08-797-689-29	Sequence 29, Appl
c 39	11	55.0	46	1	US-07-958-140-5	Sequence 5, Appl
c 40	11	55.0	46	6	PCT-US93-09166-5	Sequence 5, Appl
41	11	55.0	53	3	US-08-842-842-3	Sequence 3, Appl
42	11	55.0	79	2	US-08-447-169A-72	Sequence 72, Appl
43	11	55.0	79	3	US-08-233-012C-72	Sequence 72, Appl
44	10.8	54.0	80	4	US-08-477-527A-103	Sequence 103, App
45	10.8	54.0	80	5	US-08-481-710-103	Sequence 103, App

ALIGNMENTS

RESULT 1
US-08-850-049-18
; Sequence 18, Application US/08850049
; Patent No. 5965726
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT:
; TITLE OF INVENTION: METHOD OF ELIMINATING
; TITLE OF INVENTION: INHIBITORY/INSTABILITY REGIONS OF MRNA
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,049
; FILING DATE: 02-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050,478
; FILING DATE: 26-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/02908
; FILING DATE: 29-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/858,747
; FILING DATE: 27-MAR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MORRY, MARY J.
; REGISTRATION NUMBER: 34,398
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 BASE PAIRS
; TYPE: NUCLEIC ACID

```

; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
US-08-850-049-18

Query Match          60.0%; Score 12: DB 4; Length 50;
Best Local Similarity 63.2%; Pred. No. 3e+02;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy      2 cngaygcncarcntaygt 20
      1 | | | | | | | | | |
Db      17 CTGAAGGACAGCCATACGT 35

RESULT 2
US-08-050-478-18
; Sequence 18, Application US/08050478
; Patent No. 5972596
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT:
; APPLICANT:
; TITLE OF INVENTION: METHOD OF ELIMINATING
; TITLE OF INVENTION: INHIBITORY/INSTABILITY REGIONS OF mRNA
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/050,478
; FILING DATE: 26-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/02908
; FILING DATE: 29-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/858,747
; FILING DATE: 27-MAR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MORRY, MARY J.
; REGISTRATION NUMBER: 34,398
; REFERENCE/DOCKET NUMBER: 2026-4006US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
US-08-050-478-18

Query Match          60.0%; Score 12: DB 4; Length 50;
Best Local Similarity 63.2%; Pred. No. 3e+02;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy      2 cngaygcncarcntaygt 20
      1 | | | | | | | | | |
Db      17 CTGAAGGACAGCCATACGT 35
```

```

RESULT 3
US-08-469-318-99
; Sequence 99, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
US-08-469-318-99

Query Match          59.0%; Score 11.8; DB 5; Length 68;
Best Local Similarity 66.7%; Pred. No. 4e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      6 ygcncarcncntaygt 20
      : | | | | | | | | | |
Db      53 TCGGACGACCCCTACGT 67

RESULT 4
US-08-469-318-100/c
; Sequence 100, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
US-08-469-318-100
```

Query Match 59.0%; Score 11.8; DB 5; Length 68;
Best Local Similarity 66.7%; Pred. No. 4e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 ygcncarcctaygt 20
:|||:|||||:
Db 20 TGCGCAGCCCTACT 6

RESULT 5
PCT-US95-01185-99

; Sequence 99, Application PC/TUS9501185
; GENERAL INFORMATION:

; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01185

; FILING DATE: 02-FEB-1995
; CLASSIFICATION:

; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/192325

; FILING DATE: 14-FEB-1994
; INFORMATION FOR SEQ ID NO: 99:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid

; DESCRIPTION: /desc = "synthetic DNA"
PCT-US95-01185-99

Query Match 59.0%; Score 11.8; DB 6; Length 68;
Best Local Similarity 66.7%; Pred. No. 4e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 ygcncarcctaygt 20
:|||:|||||:
Db 53 TGCGCAGCCCTACT 67

RESULT 6
PCT-US95-01185-100/c

; Sequence 100, Application PC/TUS9501185
; GENERAL INFORMATION:

; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01185

; FILING DATE: 02-FEB-1995
; CLASSIFICATION:

; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/192325

; FILING DATE: 14-FEB-1994
; INFORMATION FOR SEQ ID NO: 100:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
PCT-US95-01185-100

Query Match 59.0%; Score 11.8; DB 6; Length 68;
Best Local Similarity 66.7%; Pred. No. 4e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 ygcncarcctaygt 20
:|||:|||||:
Db 20 TGCGCAGCCCTACT 6

RESULT 7
US-07-734-225A-37

; Sequence 37, Application US/07734225A
; Patent No. 5320840
; GENERAL INFORMATION:

; APPLICANT: Camble, Roger
; APPLICANT: Timms, David
; APPLICANT: Wilkison, Anthony J
; TITLE OF INVENTION: CONTINUOUS RELEASE PHARMACEUTICAL
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DABRY & CUSHMAN
; STREET: Eleventh Floor, 1615 L. Street, N.W.
; CITY: Washington.
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/734, 225A
; FILING DATE: 19910722
; CLASSIFICATION: 514

; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9016138.1
; FILING DATE: 23-JUL-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9018414.4
; FILING DATE: 23-AUG-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9018415.1
; FILING DATE: 23-AUG-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9018416.9
; FILING DATE: 23-AUG-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9018417.1
; FILING DATE: 23-AUG-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9018418.5
; FILING DATE: 23-AUG-1990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)861-3000
; TELEFAX: (202)822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 37:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-734-225A-37

Patent No. 5766593
GENERAL INFORMATION:
APPLICANT: Lichenstein, Henri S.
APPLICANT: Wright, Samuel D.
TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 PEPTIDES
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN INC.
STREET: 1840 Denavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1720
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/366,953A
FILING DATE: 30-DEC-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Cook Ph.D., Robert R.
REGISTRATION NUMBER: 31,602
REFERENCE/DOCKET NUMBER: A-324
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-366-953A-18

Query Match 58.0%; Score 11.6; DB 2; Length 33;
Best Local Similarity 57.9%; Pred. No. 4.9e+02;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 gcgagcncarcntayg 19
|||:|:|:|:|:|:|:|:|:|:
Db 10 gccgacgcgcgcagatag 28

RESULT 11
US-08-472-255A-89/c
Sequence 89, Application US/08472255A
Patent No. 576653
GENERAL INFORMATION:
APPLICANT: PARMA, DAVID
APPLICANT: GOLD, LARRY
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS
TITLE OF INVENTION: NO SELECTINS (AS AMENDED)
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
COMPUTER: IBM pc compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,255A
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131

FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX40-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All C's are 2'-NH2 cytosine
FEATURE:
OTHER INFORMATION: All U's are 2'-NH2 uracil
US-08-472-255A-89

Query Match 58.0%; Score 11.6; DB 2; Length 71;
Best Local Similarity 57.9%; Pred. No. 5.2e+02;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 cngaycncarcntaygt 20
|:|:|:|:|:|:|:|:|:|:
Db 36 CAGACGCGCTACTTATGT 18

RESULT 12
US-08-479-724A-89/c
Sequence 89, Application US/08479724A
Patent No. 5780228
GENERAL INFORMATION:
APPLICANT: PARMA, DAVID
APPLICANT: GOLD, LARRY
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS
TITLE OF INVENTION: TO LECTINS
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
COMPUTER: IBM pc compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,724A
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
ATTORNEY/AGENT INFORMATION:

NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX40-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All C's are 2'-NH2 cytosine
FEATURE:
OTHER INFORMATION: All U's are 2'-NH2 uracil
US-08-479-724A-89

Query Match 58.0%; Score 11.6; DB 2; Length 71;
Best Local Similarity 57.9%; Pred. No. 5.2e+02;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 cngaygcncarcntaygt 20
| ||||| : | ||||
DB 36 CAGACGCGCTACTTTATGT 18

RESULT 13
US-08-472-256B-89/C
Sequence 89, Application US/08472256B
Patent No. 6001988
GENERAL INFORMATION:
APPLICANT: PARMA, DAVID
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS
TITLE OF INVENTION: TO LECTINS
NUMBER OF SEQUENCES: 177
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
COMPUTER: IBM pc compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,256B
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 538
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:

LENGTH: 71 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All C's are 2'-NH2 cytosine
FEATURE:
OTHER INFORMATION: All U's are 2'-NH2 uracil
US-08-472-256B-89

Query Match 58.0%; Score 11.6; DB 5; Length 71;
Best Local Similarity 57.9%; Pred. No. 5.2e+02;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 cngaygcncarcntaygt 20
| ||||| : | ||||
DB 36 CAGACGCGCTACTTTATGT 18

RESULT 14
PCT-US96-09455A-89/C
Sequence 89, Application PC/TUS9609455A
GENERAL INFORMATION:
APPLICANT: PARMA, et al.
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID
TITLE OF INVENTION: LIGANDS TO LECTINS
NUMBER OF SEQUENCES: 390
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
COMPUTER: IBM pc compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09455A
FILING DATE: 05 JUNE 1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,724
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/472,256
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/472,255
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,829
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX40C/PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All C's are 2'-NH2 cytosine

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:28 ; Search time 1236.38 Seconds

(without alignments)
-15.736 Million cell updates/sec

Title: US-09-164-714-5

Perfect score: 20

Sequence: 1 atnccrtangcngtrtlytt 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 882769 segs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: GenBml:*
2: gb_ba1:*
3: gb_ba2:*
4: gb_om:*
5: gb_ov:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_hg1:*
33: gb_hg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_hg3:*
42: gb_hg4:*
43: gb_hg5:*
44: gb_hg6:*

45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	12.6	63.0	91	3 BTU19470	U19470 Bos taurus
2	12.6	63.0	91	3 BTU19472	U19472 Bos taurus
3	12.6	63.0	91	3 CDU19479	U19479 Dama dama 1
4	12.6	63.0	91	3 CHU19477	U19477 Capra hircu
5	12.6	63.0	91	3 OAU19474	U19474 Ovis aries
6	12.6	63.0	91	3 OAU19475	U19475 Ovis aries
7	12.6	63.0	97	5 AR030732	AR030732 Sequence
8	12.2	61.0	48	5 113041	113041 Sequence 18
9	12.2	61.0	77	5 140687	140687 Sequence 18
10	12	60.0	49	5 AR061917	AR061917 Sequence
11	12	60.0	49	5 E06680	E06680 DNA probe t
12	12	60.0	65	5 A87799	A87799 Sequence 18
13	12	60.0	65	5 A87800	A87800 Sequence 19
14	11.8	59.0	61	5 E17241	E17241 Primer. 7/1
15	11.8	59.0	87	5 AR003822	AR003822 Sequence
16	11.8	59.0	87	5 AR010158	AR010158 Sequence
17	11.8	59.0	87	5 AR034690	AR034690 Sequence
18	11.8	59.0	87	5 AR048571	AR048571 Sequence
19	11.8	59.0	87	5 AR055364	AR055364 Sequence
20	11.8	59.0	87	5 140575	140575 Sequence 13
21	11.6	58.0	44	5 113034	113034 Sequence 11
22	11.6	58.0	48	5 113042	113042 Sequence 19
23	11.6	58.0	55	7 YSAMTEO3SB	D16219 Candida alb
24	11.6	58.0	91	3 BTU19471	U19471 Bos taurus
25	11.4	57.0	29	5 144871	144871 Sequence 8
26	11.4	57.0	40	5 AR029226	AR029226 Sequence
27	11.4	57.0	80	9 HSU14601	U14601 Homo sapien
28	11.4	57.0	80	9 HUMSINE	L16226 Human SINE
29	11.2	56.0	39	5 AR031240	AR031240 Sequence
30	11.2	56.0	48	5 AR031254	AR031254 Sequence
31	11.2	56.0	81	16 AF040873	AF040873 Hepatitis
32	11.2	56.0	91	12 AB010315	AB010315 Mus muscu
33	11.1	55.0	27	5 140766	140766 Sequence 97
34	11.1	55.0	29	5 AR024157	AR024157 Sequence
35	11.1	55.0	33	5 A69052	A69052 Sequence 10
36	11.1	55.0	60	8 CNS01A9T	A112665 Botrytis
37	11.1	55.0	60	12 S63846S1	S63846 dopamine D3
38	11.1	55.0	75	10 HSNCDX10D	Z27430 H.sapiens N
39	11.1	55.0	78	10 HSA243199	AJ243199 Homo sapi
40	11.1	55.0	83	4 XU72860	U72860 Xenopus lae
41	11.1	55.0	91	3 CDU19480	U19480 Dama dama 1
42	11.1	55.0	91	3 CHU19476	U19476 Capra hircu
43	11.1	55.0	91	3 CHU19478	U19478 Capra hircu
44	11.1	55.0	91	3 OAU19473	U19473 Ovis aries
45	10.8	54.0	20	5 E08796	E08796 PCR primer

ALIGNMENTS

[illegible]

JOURNAL	Direct Submission
TITLE	Submitted (04-JAN-1995) David M. Irwin, Clinical Biochemistry, University of Toronto, 100 College St., Toronto, Ont M5G 1L5, Canada
FEATURES	Location/Qualifiers 1..91 /organism="Bos taurus" /isolate="K4" /db_xref="taxon:9913" exon <1..>91 /number=2 <1..>91 codon_start=2 product="lysozyme" protein_id="AA85546.1" /db_xref="GI:84120" translation="WESNYNTRATNNYNGDKSTDYGIFQINSRW"
CDS	
BASE COUNT	36 a 17 c 18 g 20 t
ORIGIN	
Query Match	63.0%; Score 12.6; DB 3; Length 91; Best Local Similarity 60.0%; Pred. No. 1.2e+04; Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Oy	1 atnccrtangcngtttctt 20 :: :: :: Db 69 ATCCCATATACAGTCTTTT 50
RESULT 3	
LOCUS	CDU19479/c 91 bp DNA MM 24-JAN-1996
DEFINITION	Dama dama lysozyme gene, isolate K2, partial exon 2 and partial cds.
ACCESSION	U19479
VERSION	U19479.1 GI:841233
KEYWORDS	
SOURCE	. fallow deer. Cervus dama
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae; Cervidae; Cervinae; Cervus.
REFERENCE	1 (bases 1 to 91) Irwin,D.M. Evolution of the bovine lysozyme gene family: changes in gene expression and reversion of function J Mol. Evol. 41 (3), 299-312 (1995)
AUTHORS	JOURNAL MEDLINE REFERENCE AUTHORS TITLE JOURNAL
FEATURES	Location/Qualifiers 1..91 /organism="Cervus dama" /isolate="K2" /db_xref="taxon:30532" exon <1..>91 /number=2 <1..>91 codon_start=2 product="lysozyme" protein_id="AA85553.1" /db_xref="GI:84134" translation="WESNYWQATNNYNGDKSTDYGIFQINSRW"
CDS	
BASE COUNT	37 a 19 c 18 g 17 t
ORIGIN	
Query Match	63.0%; Score 12.6; DB 3; Length 91;

ACCESSION	A87800
VERSION	A87800.1
KEYWORDS	GI:67369391
SOURCE	.
ORGANISM	human.
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 65)
AUTHORS	Karayam, U., and Hong, S.-S.
TITLE	USE OF A POLYPEPTIDE AS CELL RECEPTOR FOR ADENOVIRUSES
JOURNAL	Patent: WO 9833929-A 06-AUG-1998; KARAYAN LUCIE (FR); HONG SAW SEE (FR) Location/Qualifiers
FEATURES	location/Qualifiers
source	1..65 /organism="Homo sapiens" /isolate="OLIGONUCLEOTIDE DE SYNTHÈSE CODANT POUR AA 157-176 DU MHC-I"
BASE COUNT	/db_xref="taxon:9606" 20 a 27 c 6 g 12 t
ORIGIN	
Query Match	60.0%; Score 12; DB 5; Length 65;
Best Local Similarity	63.2%; Pred. No. 2,6e+04;
Matches 12:	Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY	1 atnccrtangcngtcttyt 19 Db 38 ATGGCTTAGCGGATATT 20
RESULT 14	E17241/c
LOCUS	E17241 61 bp DNA PAT 28-JUL-1999
DEFINITION	Primer.
ACCESION	E17241
VERSION	E17241.1 GI:5711924
KEYWORDS	JP 1998257893-A/18.
SOURCE	unidentified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 61)
AUTHORS	Nakamura,K. and Hanai,N.
TITLE	HOMAN COMPLEMENTARITY DETERMINING REGION(CDR) TRANSPLANTED ANTIBODY AGAINST GANGLIOSIDE GM2
JOURNAL	Patent: JP 1998257893-A 29-SEP-1998; KYOMA HAKKO KOGYO CO LTD
COMMENT	OS None OC Artificial sequences. PN JP 1998257893-A/18 PD 29-SEP-1998 PF 19-MAR-1997 JP 1997066981 PI NAKAMURA KAZUYASU, HANAII NOBUO PC C12N15/09,A6IK39/395,A6IK39/395,C07K16/18,C12N5/10,C12P21/08, PC G01M33/531,G01N33/574,G01N33/577,(C12N5/10,C12R1:91), PC (C12P21/08,C12R1:91); CC strandedness: Single; CC topology: Linear; CC hypothetical: No; FH key Location/Qualifiers FT source 1..61 /organism='Artificial sequences'. FT location/Qualifiers 1..61 /organism="unidentified" /db_xref="taxon:32644"
BASE COUNT	18 a 24 c 14 g 5 t
ORIGIN	
Query Match	59.0%; Score 11.8; DB 5; Length 61;

Best Local Similarity 66.7%; Pred. No. 3.3e+04;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY	6 rtangcngtrtlytt	20
----	-------------------	----

Db 26 GTAGGCTGTGTTCTT 12

RESULT 15
AR003822/C
10000000 07 11 00000000

LOCUS	AR003822	87 bp	DNA	PAT	04-DEC-1998
-------	----------	-------	-----	-----	-------------

DEFINITION	Sequence 135 from patent US 5744580.
ACCESSION	AB002823

ACCESSION	AR003822
VERSION	AR003822 1
	GT:3965081

VERSION 01:000001
KEYWORDS ,

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 87)

REFERENCE
1 (bases 1 to 8)
AUTHORS
Better, M.D., Carroll, S.F., and Studnicka, G.M.

TITLE Immunotoxins comprising ribosome-inactivating proteins

JOURNAL Patent: US 5744580-A 135 28-APR-1998;

FEATURES	Location/Qualifiers
Source	107

```
source
1.:8/
/organisms="unknown"
```

	BASE COUNT	ORGANISM - UNKNOWN
22 a	22 c	19 q
		24 t

ORIGIN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
--------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

Query Match EQ 08: EQ010 11 8: DB 5: T0004h 07:

Query Match	39.06;	Score 11.8;	DB 3;	Length 8/;
Best Local Similarity	66.78;	Pred. NO. 3.3e+04;		

```

Score: 200.00; E-Value: 1.2e-06; Identity: 98.7%; Positives: 100%;
Matches: 10; Conservative: 3; Mismatches: 2; Indels: 0; Gaps: 0

```

QY 6 rtangcngtrtlytt 20

Db 44 ATAGGCAGTGTTCCT 30
:| | | :| :| |

```
Search completed: June  4, 2000, 16:05:30
Job time: 27882 sec
```

Job time: 27882 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:01 ; Search time 322.35 Seconds
(without alignments)
15.523 Million cell updates/sec

Title: US-09-164-714-5

Perfect score: 20

Sequence: 1 atnccrtaangcngtrtytt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12.6	63.0	97	1	V09676 Human cathepsin K
2	12.2	61.0	48	1	Oligo probe TAP160
3	12.2	61.0	77	1	Secretory phosphol
4	12.2	60.0	49	1	Nucleic acid probe
5	12.2	60.0	49	1	u. urealyticum 16S
6	12.2	60.0	59	1	Synthetic Fn3 gene
7	11.8	59.0	61	1	Primer for coding
8	11.8	59.0	80	1	De-immunised 708 V
9	11.8	59.0	87	1	HUN-G15. Antibodie
10	11.8	59.0	87	1	Humanised H65 heav
11	11.6	58.0	44	1	TAP probe (1). New
12	11.6	58.0	44	1	Probe for bovine t
13	11.6	58.0	48	1	Oligo probe TAP113
14	11.4	57.0	29	1	Salmonella enterit
15	11.4	57.0	34	1	BH3 interacting do
16	11.4	57.0	40	1	RF-X polynucleotid
17	11.4	57.0	58	1	Human gene signatu
18	11.2	56.0	21	1	Primer R1 for H.py
19	11.2	56.0	62	1	L-selectin orphan
20	11.2	56.0	67	1	Staphylococcus aur
21	11.2	55.0	23	1	nachr beta2 subuni
22	11.1	55.0	27	1	Secretory phosphol
23	11.1	55.0	29	1	Herpes simplex vir
24	11.1	55.0	30	1	Oligonucleotide ZC
25	11.1	55.0	31	1	Nucleotide fragmen
26	11.1	55.0	33	1	Primer AB050 for C
27	11.1	55.0	51	1	Murine osteoproteg
28	11.1	55.0	53	1	Staphylococcus aur
29	11.1	55.0	62	1	Downstream sequenc
30	11.1	55.0	88	1	Human gene signatu
31	10.8	54.0	100	1	Downstream sequenc
32	10.8	54.0	20	1	Lactobacillus sp.
33	10.8	54.0	31	1	Human biallelic po
34	10.8	54.0	68	1	Embryonic DNA sequ

35	10.8	54.0	79	1	003346	Sequence from chro
36	10.6	53.0	19	1	X07468	Human BS124 specif
37	10.6	53.0	22	1	Q82111	Chromosome 11 (loc
38	10.6	53.0	24	1	T10614	LN clone E2 Vbeta8
39	10.6	53.0	26	1	V41471	Nucleotide sequenc
40	10.6	53.0	29	1	Q76064	hpy Primer set 13,
41	10.6	53.0	31	1	Q84448	Mycobacterium marl
42	10.6	53.0	31	1	V15800	Primer for wild ty
43	10.6	53.0	31	1	X06424	Human biallelic po
44	10.6	53.0	32	1	T30768	Yeast hls3 gene PC
45	10.6	53.0	40	1	V08300	Primer for HIV-1 p

ALIGNMENTS

RESULT 1
V09676/c
ID V09676 standard: DNA; 97 BP.
AC V09676;
DT 20-JUL-1998 (first entry)
DE Human cathepsin K gene exon 8.
KW Cathepsin K; human; osteoporosis; periodontal disease;
KW Paget's disease; Gaucher's disease; Alzheimer's disease;
KW central nervous system inflammation; hyperparathyroidism;
KW bone degradation; dental implant degradation; metastasis; tumour;
KW diagnosis; therapy; ss.
OS Homo sapiens.
PN EP-812916-A2.
PD 17-DEC-1997.
PF 19-MAY-1997; 303395.
PR 26-AUG-1996; US-026083.
PR 14-JUN-1996; US-019942.
PR 17-JUN-1996; US-020273.
PR 26-AUG-1996; WO-014026.
PA (HUMA-) HUMAN GENOME SCI INC.
PA (GENO-) INST GENOMIC RES.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PI Adams MD, Blake JA, Dabouck CM, Drake FH, Fitzgerald LM,
PI Fraser CM, Gowen M, Hastings GA, Kirkness EF, Lee NH,
PI Riord J;
PI WPI: 98-034977/04.
DR P-PSDB: M39270.
PT DNA encoding human cathepsin K - useful for diagnosing and treating
PT diseases associated with cathepsin K e.g. osteoporosis, bone
PT degradation, metastatic tumours, etc
PS Claim 13; Page 45; 84pp; English.
CC This genomic DNA sequence comprises the exon 8 of the human
CC cathepsin K gene (see V09660). It encodes a 32-amino acid
CC C-terminal peptide (see W39270) of cathepsin K (see W39216).
CC Polymorphisms in the 8 exons of the cathepsin K gene serve as
CC markers for osteoporosis, periodontal disease, Paget's disease,
CC Gaucher's disease, central nervous system inflammation, Alzheimer's
CC disease, hyperparathyroidism, bone degradation, metastatic tumours,
CC and degradation of bone implants and prostheses, especially dental
CC implants. Ribozymes based e.g. on splice junctions of cathepsin K
CC gene polynucleotides can be used to inhibit cathepsin K gene
CC expression. Isolated polypeptides encoded by exons 1, 2, 3, 4, 5,
CC 6, 7 and 8 are claimed (see W39264-70), as well as those encoded by
CC exon pairs 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 2-4, 2-5, 2-6, 2-7, 2-8,
CC 3-4, 3-5, 3-6, 3-7, 3-8, 3-6, 4-5, 4-6, 4-7, 4-8, 5-7, 5-8 and 6-8.
CC They can be used to treat a patient in need of cathepsin K or to
CC identify compounds which bind to and inhibit activation of
CC cathepsin K.
SQ Sequence 97 BP; 31 A; 24 C; 25 G; 17 T;
Query Match 63.0%; Score 12.6; DB 1; Length 97;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
1 atnccrtaangcngtrtytt 20
|||:||||:|

DB 69 ATGCCACAGCGCTTCTT 50

RESULT 2

079194/c

ID 079194 standard; DNA; 48 BP.

AC 079194;

DT 04-MAY-1995 (first entry)

DE Oligo probe TAP160a for bovine tracheal antimicrobial peptide gene.

CC Antimicrobial peptide; TAP; probe; ss.

OS Synthetic.

PN WO9426106-A.

PD 24-NOV-1994.

PF 11-MAY-1994; U05257.

PR 11-MAY-1993; US-060822.

PA (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.

PI Bevins CL, Diamond G, Zasloff MA.

DR WPI. 95-006233/01.

PT New antimicrobial peptide and precursor from mammalian trachea -

PT and related DNA, vector and transformed cells, useful as

PT Pharmaceuticals and disinfectants

PT Example; page 27; 69 pp; English.

CC A mammalian TAP (R66203) was isolated by a combination of

CC size-exclusion, ion-exchange, and reverse-phase chromatographic

CC fractionations using antimicrobial activity against a strain of

CC E.coli as a functional assay. The mol. wt. of TAP is 4085 Da. The

CC predicted PI is 13.0 and there are no aromatic residues. A cDNA

CC sequence corresp. to the precursor peptide was cloned (see 079533/

CC R66204) and the genomic TAP gene (079534/R66205) was isolated.

CC PCR primers for amplifying the TAP gene from bovine genomic DNA

CC were TAPG5 and TAPG10a. Oligo probes for hybridisation were

CC TAP48a, TAP54s, TAP160a, TAP13S, TAP18S.

CC Sequence 48 BP; 14 A; 14 C; 8 G; 12 T;

Query Match 61.0%; Score 12.2; DB 1; Length 48;

Best Local Similarity 61.1%; Pred. No. 3.3e+02;

Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 atnccrtangcngtrttt 18

DB 47 ATCCCTAGAGCTGTCTT 30

RESULT 3

T88489

ID T88489 standard; RNA; 77 BP.

AC T88489;

DT 17-DEC-1997 (first entry)

DE Secretory phospholipase A2 nucleic acid ligand used in diagnosis.

CC Nucleic acid ligand; secretory phospholipase A2; inflammation;

CC anti-inflammatory; rheumatoid arthritis; septic shock; VEGF;

CC vascular endothelial growth factor; angiogenesis; psoriasis;

CC metastasis; neoplasia; haemangioma; human immunodeficiency virus;

CC HIV; diagnosis; treatment; specific; high affinity; ss.

OS Synthetic.

PN WO9627604-A1.

PD 12-SEP-1996.

PF 04-MAR-1996; U02942.

PR 19-MAY-1995; US-447172.

PR 06-MAR-1995; US-399412.

PR 19-MAY-1995; US-447169.

PA (NEXS-) NEXSTAR PHARM INC.

PI Gold L, Janjic N, Lochrie M, Parma D;

DR WPI. 96-425373/42.

PT New non-naturally occurring nucleic acid ligands to secretory

PT phospholipase A2, VEGF and HIV-1 gag - useful, e.g., for diagnosis

PT and treatment of inflammation, angiogenesis and HIV infection.

PS Claim 9; Page 91; 169pp; English.

CC T88481-T88593 are nucleic acid ligands (NAL) for secretory phospholipase

CC A2 (sPLA2). Other non-naturally occurring ligands for vascular

CC endothelial growth factor (VEGF) and HIV-1 gag were also made. The NAL

CC were identified using the SELEX procedure. They are used as diagnostic

CC or treatment agents. For example, those for sPLA2 are used to study/
CC diagnose inflammation and as anti-inflammatory agents e.g. for the
CC treatment of rheumatoid arthritis and septic shock; those for VEGF are
CC used to treat angiogenesis-mediated diseases e.g. psoriasis, haemangioma,
CC metastasis; and those for gag are used to inhibit HIV-1 replication.
CC NAL can also generally be used to identify the corresponding target.
CC Sequence 77 BP; 22 A; 19 C; 19 G; 17 U;

Query Match 61.0%; Score 12.2; DB 1; Length 77;

Best Local Similarity 27.8%; Pred. No. 3.5e+02;

Matches 5; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY 2 tncctangcngtrttt 19

DB 35 UNCCGAGACGCAUUCU 52

RESULT 4

O54447

ID O54447 standard; DNA; 49 BP.

AC O54447;

DT 04-JUL-1994 (first entry)

DE Nucleic acid probe for Ureaplasma urealyticum.

CC Probe; screening; detection; identification; septic abortion;

CC nongonococcal urethritis; pelvic inflammatory disease;

CC postpartum fever; clinical sample; swab; lavage;

CC Mycoplasma genitalium; Mycoplasma hominis; Ureaplasma urealyticum;

OS Synthetic.

PN EP-576742-A.

PD 05-JAN-1994.

PF 04-JUN-1992; 305125.

PR 04-JUN-1992; EP-305125.

PA (STAD) AMOCO CORP.

PI Pelletier DA, Weisburg WG;

DR WPI; 94-009116/02.

PT Nucleic acid probes for the detection of genital mycoplasma - 1s

PT able to hybridise with RNA or rDNA of pathogenic mycoplasma

PS Claim 9; Page 10; 23pp; English.

CC Probes specific for Ureaplasma urealyticum (O54446-52),

CC Mycoplasma genitalium (O54453-58) and Mycoplasma hominis (O63723-25)

CC can be used for the selection of these three presumptive

CC etiological agents of nongonococcal urethritis, septic abortion,

CC inflammatory disease and postpartum fever from clinical samples such

CC as genital swabs, genital lavage, sputum, throat swabs, blood, urine,

CC cerebrospinal fluid, skin, biopsy, saliva, synovial fluid, bronchial

CC wash, bronchial lavage or other tissue or fluid samples from humans

CC or animals. The probes can also be used to confirm the presence of

CC the genital mycoplasmas in liquid or semi-solid in vitro culture.

CC This probe is specific for the 16S rRNA of Ureaplasma urealyticum.

CC Sequence 49 BP; 11 A; 13 C; 8 G; 17 T;

Query Match 60.0%; Score 12; DB 1; Length 49;

Best Local Similarity 70.6%; Pred. No. 4.3e+02;

Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 atnccrtangcngtrttt 17

DB 33 ATGCCCTATGCGGTATT 49

RESULT 5

V64974

ID V64974 standard; DNA; 49 BP.

AC V64974;

DT 05-FEB-1999 (first entry)

DE U. urealyticum 16S ribosomal RNA probe 2256.

CC 16S ribosomal RNA; rRNA; probe; detection; etiological agent; PID;

CC nongonococcal urethritis; pelvic inflammatory disease; salpingitis;

CC genital infection; ss.

OS Synthetic.


```
OS Ureaplasma urealyticum.
PN US843667-A.
PD 01-DEC-1998.
PF 22-MAR-1998.
PR 22-MAR-1991; 673661.
PA (STAD ) AMOCO CORP.
PI Pelletier DA, Weisburg WC;
DR WPI: 99-044571/04.
PT Nucleic acid hybridisation probes - for detecting genital mycoplasma
PS species
PT Claim 3; Column 19-20; 16pp; English.
CC V64973-V64989 are probes used in the detection of 16S ribosomal RNA or
CC DNA (rRNA, rDNA) regions from mycoplasma etiological agents of
CC nongonococcal urethritis, pelvic inflammatory disease (PID), salpingitis
CC and other genital infections. These probes can be used in the detection
CC of Mycoplasma hominis, Ureaplasma urealyticum or Mycoplasma genitalium
CC in a sample.
SQ Sequence 49 BP; 11 A; 13 C; 8 G; 17 T;

Query Match 60.0%; Score 12; DB 1; Length 49;
Best Local Similarity 70.6%; Pred. No. 4.3e+02;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 atnccrtangncgttrtt 17
Db 33 ATGCCCTATGCGGTAAT 49

RESULT 6
X05501
ID X05501 standard; DNA; 59 BP.
AC X05501;
DT 14-APR-1999 (first entry)
DE Synthetic Fn3 gene constructing oligo FN1F.
KM Fibronectin type III; Fn3; monobody; beta-strand domain; loop region;
KW specific binding partner; SBP; catalysts; LRS; ss.
OS Synthetic.
OS Homo sapiens.
PN W09856915-A2.
PD 17-DEC-1998.
PF 12-JUN-1998; U12099.
PR 12-JUN-1997; US-049410.
PA (RESE ) RESEARCH CORP TECHNOLOGIES INC.
PI Koida S;
DR WPI: 99-060331/05.
PT Production of antibody compounds, particularly catalytic antibodies
PT - using a fibronectin type III molecular scaffolding comprising
PT beta-strand domain sequences and modified in one or more loop
PT sequences
PS Example 1; Page 23; 96pp; English.
CC The invention relates to a synthetic fibronectin type III (Fn3)
CC polypeptide monobody that comprises Fn3 beta-strand domain sequences
CC that are linked to loop region sequences (LRSs). One or more of the loop
CC sequences in the synthetic Fn3 vary by deletion, insertion, or
CC replacement of at least 2 amino acids from the corresponding LRSs in
CC wild-type Fn3. Host cells containing an expression vector comprising the
CC synthetic Fn3 nucleic acid are used for the production of the Fn3
CC monobody. The invention also provides methods of identifying the amino
CC acid sequence of a polypeptide molecule (1) capable of binding to a
CC -specific binding partner (SBP) so as to form a polypeptide-SBP complex;
CC (11) capable of catalysing a chemical reaction with a catalysed rate
CC constant, Kcat, and an uncatalysed rate constant, Kuncat, such that the
CC ratio of the Kcat/Kuncat is greater than 10. Sequences X05501-18
CC represent oligos used for constructing a synthetic Fn3 gene.
SQ Sequence 59 BP; 9 A; 15 C; 18 G; 17 T;
```

```
Query Match 60.0%; Score 12; DB 1; Length 59;
Best Local Similarity 63.2%; Pred. No. 4.4e+02;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 atnccrtangncgttrtt 19
```

```
Db 5 ATCCCATATGACAGTTCT 23

RESULT 7
V08052/c
ID V08052 standard; DNA; 61 BP.
AC V08052;
DT 22-JAN-1999 (first entry)
DE Primer for coding sequence of ganglioside GM2 targeting antibody.
KM Ganglioside GM2; antibody; complementarily determining region; cancer;
KW anti-tumour agent; PCR primer; ss.
OS Synthetic.
OS Homo sapiens.
PN J10257893-A.
PD 29-SEP-1998.
PF 19-MAR-1997; 066981.
PR 19-MAR-1997; JP-066981.
PA (KYOW ) KYOWA HAKKO KOGYO KK.
DR WPI: 98-575904/49.
PT A human type complementarily determining region transplanted
PT antibody against ganglioside GM2 - useful as an anti-tumour agent
PT and as a diagnostic for related cancers
PS Example 3; Page 34; 66pp; Japanese.
CC This sequence represents a PCR primer for DNA encoding the antibody of
CC the invention. The antibody is a human complementarily determining
CC region transplanted antibody that reacts specifically with ganglioside
CC GM2. DNA encoding the antibody, and vectors and transformants containing
CC it, can be used for the recombinant production of the antibody. The
CC antibody itself can be used as an anti-tumour agent or as a diagnostic
CC tool for related cancers. The antibody has antitumour activity against
CC ganglioside GM2 positive cells.
SQ Sequence 61 BP; 18 A; 24 C; 14 G; 5 T;

Query Match 59.0%; Score 11.8; DB 1; Length 61;
Best Local Similarity 66.7%; Pred. No. 5.7e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 rtangncgttrtt 20
Db 26 GTAGGCTGTGTTCTT 12

RESULT 8
V81050/c
ID V81050 standard; DNA; 80 BP.
AC V81050;
DT 03-MAR-1999 (first entry)
DE De-immunised 708 Vh constructing long oligo DIVH4.
KM Non-immunogenic; epitope; T-cell; immunogenicity; immune system; SK;
KW immunoglobulin; therapeutic; streptokinase; de-immunised; 708; ss.
OS Synthetic.
PN W09852976-A1.
PD 26-NOV-1998.
PF 21-MAY-1998; G01473.
PR 14-APR-1998; GB-007751.
PR 21-MAY-1997; GB-010480.
PR 31-JUL-1997; GB-016197.
PR 28-NOV-1997; GB-025270.
PR 02-DEC-1997; US-067235.
PA (BIOV-) BIOVATION LTD.
PI Carr FJ;
DR WPI: 99-045301/04.
PT Reducing immunogenicity of proteins - by modifying the amino acid
PT sequence of the protein to eliminate potential epitopes for T-cells
PT of a given species
PS Example 3; Fig 16; 77pp; English.
CC The invention relates to a method for the production of non-immunogenic
CC proteins. The method comprises determining at least part of the amino
CC acid sequence of the protein; (b) identifying in the amino acid sequence
CC one or more potential epitopes for T-cells (T-cell epitopes) of the given
CC species; and (c) modifying the amino acid sequence to eliminate at least
```

CC one of the T-cell epitopes identified in step (b) thereby to eliminate or
CC reduce the immunogenicity of the protein when exposed to the immune
CC system of the given species. A method of analysing a pre-existing protein
CC to predict the basis for immunogenic responses is also provided. The
CC methods can be used particularly for reducing the immunogenicity of
CC immunoglobulins or therapeutic proteins, e.g. streptokinase (SK). The
CC products can be used for diagnosis and therapy. Sequences V81047-68
CC represent oligonucleotides used for the construction of de-immunised 708
CC Vh and Vk.
SQ Sequence 80 BP; 24 A; 19 C; 21 G; 16 T;

Query Match 59.0%; Score 11.8; DB 1; Length 80;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 6 rtangengrttytt 20
:|||:|||:||||
Db 55 GTRAGGCTGTCTT 41

RESULT 9
044074/c
ID 044074 standard; DNA; 87 BP.
AC 044074;
DT 28-OCT-1993 (first entry)
DE HUH-G15.
KW Antibody; variable domain; light; L; heavy; H; PCR;
KW polymerase chain reaction; H65; monoclonal antibody; Mab; ss.
OS Synthetic.
PN WO9311794-A.
PD 24-JUN-1993.
PR 14-DEC-1992; 010906.
PR 13-DEC-1991; US-808464.
PA (XOMA) XOMA CORP.
PI Fishwild DM, Kohn FR, Little RG, Studnicka GM;
DR WPI: 93-213827/26.
PT Antibodies prepn. used for treatment of auto-immune diseases - by
PT replacement of critical residues to reduce immunogenicity but
PT retain binding affinity, etc.
PS Example 6: Fig 7B; 160pp: English.
CC Example 6 describes the construction of the gene encoding the
CC humanised H65 heavy chain contg. the moderate risk residues.
CC The humanised heavy chain was assembled from the oligonucleotides
CC given in 044070-75.
SQ Sequence 87 BP; 22 A; 22 C; 19 G; 24 T;

Query Match 59.0%; Score 11.8; DB 1; Length 87;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 6 rtangengrttytt 20
:|||:|||:||||
Db 44 ATRAGCAGTGTCTT 30

RESULT 10
V31078/c
ID V31078 standard; DNA; 87 BP.
AC V31078;
DT 18-AUG-1998 (first entry)
DE Humanised H65 heavy chain variable region construction oligo HUH-G15.
KW Humanised; human; mouse; CD5; anti-CD5 antibody; immunoglobulin;
KW depletion; cytotoxic; immunocjugate; fusion protein; psoriasis;
KW autoimmune disease; rheumatoid arthritis; type I diabetes; ss.
OS Synthetic.
PN Mus sp.
OS Homo sapiens.
PN US5770196-A.
PD 23-JUN-1998.
PF 07-JUN-1995; 472788.
PR 23-JUN-1995; US-082842.

PR 13-DEC-1991; US-808464.
PR 14-DEC-1992; WO-010906.
PR 07-JUN-1995; US-472788.
PA (XOMA) XOMA CORP.
PI Studnicka GM;
DR WPI: 98-376744/32.
PT Depletion of CD5-positive cells in vivo - using anti-CD5 antibodies
PT with humanised variable regions
PS Example 7; Fig 7B; 77pp: English.
CC A method has been developed of depleting CD5+ cells in an animal. The
CC method comprises administering a cytotoxic protein containing a modified
CC immunoglobulin (Ig) variable domain, where the protein is an anti-CD5
CC molecule or an immunocjugate or fusion protein containing an anti-CD5
CC Ig molecule, and where the modified Ig variable domain comprises at
CC least one of (a) a modified light chain variable region (see W58478 or
CC W58480), and (b) a modified heavy chain variable region (see W58479 or
CC W58481), where W58478 and W58479 are humanised forms of the H65 light
CC and heavy chain variable domains with low risk amino acid substitutions
CC [i.e. low risk of reducing antigen-binding specificity.] and W58480 and
CC W58481 are humanised forms of the H65 light and heavy chain variable
CC domains with moderate risk amino acid substitutions and are present in
CC humanised H65 antibody he3 (ATCC HB 11206). The method is useful for
CC treating autoimmune diseases, especially systemic lupus erythematosus,
CC rheumatoid arthritis, psoriasis or type I diabetes. The present sequence
CC represents an oligonucleotide used in the construction of humanised
SQ anti-CD5 immunoglobulin H65 heavy chain variable region DNA sequence.
Sequence 87 BP; 22 A; 22 C; 19 G; 24 T;

Query Match 59.0%; Score 11.8; DB 1; Length 87;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 6 rtangengrttytt 20
:|||:|||:||||
Db 44 ATRAGCAGTGTCTT 30

RESULT 11
025193
ID 025193 standard; DNA; 44 BP.
AC 025193;
DT 19-NOV-1992 (first entry)
DE TAP probe (1).
KW Tracheal antimicrobial peptide; respiratory mucosa; RFLP; ss.
OS Synthetic.
PN WO9207873-A.
PD 14-MAY-1992.
PF 24-OCT-1991; 007996.
PR 25-OCT-1990; US-603451.
PA (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
PI Bevins CL, Diamond G, Zasloff MA;
DR WPI: 92-183626/22.
PT New tracheal antimicrobial peptide and encoding DNA - used for
PT treating bacterial and fungal infections, used in mouthwashes,
PT contact disinfectants etc.
PS Disclosure; Page 22; 42pp: English.
CC The probes represented in 025193-94 are used in the screening
CC of bovine tracheal cDNA library.
SQ Sequence 44 BP; 12 A; 7 C; 12 G; 13 T;

Query Match 58.0%; Score 11.6; DB 1; Length 44;
Best Local Similarity 64.7%; Pred. No. 7e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 atncrtangngtrtt 17
|||:|||:|||:|
Db 2 ATCTGTGAAGCTGTGT 18

RESULT 12
079539

ID Q79539 standard; DNA; 44 BP.
AC Q79539;
DT 04-AUG-1995 (first entry)
DE Probe for bovine tracheal antimicrobial peptide.
KW Antimicrobial peptide; TAP; probe; ss.
OS Synthetic.
PN M09426106-A.
PD 24-NOV-1994.
PF 11-MAY-1994; U05257.
PR 11-MAY-1993; US-060822.
PA (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
PI Bevins CL, Diamond G, Zasloff MA;
DR WPI; 95-006233/01.
PT New antimicrobial peptide and precursor from mammalian trachea -
PT pharmaceuticals and disinfectants
PS Example; page 44; 69 pp; English.
CC In order to clone TAP cDNA, degenerate oligo primers (Q79535,
CC Q79541, Q79536) were designed corresp. to AAs 1-6 (BT-40-1
CC and -2) and 21-26 (BT-40-3) respectively. Sense primers 1
CC (Q79535) and 2 (Q79541) had the EcoRI recognition site incorporated
CC on the 5' end, and antisense primer 3 (Q79536) had the SstI
CC recognition site on its 5' end. These primers were used in a PCR
CC using bovine tracheal cDNA as template DNA and were expected to
CC amplify the nucleotide sequence coding for AAs 1-26. The principle
CC DNA product was 90 bp in length, as expected. This indicated the
CC peptide was present in the library and yielded a DNA template for
CC probe synthesis. The cDNA library from bovine tracheal epithelium
CC was screened using three different probes in parallel: the PCR
CC product; and two 'best-guess' synthetic oligo probes (Q79538,
CC Q79539). The sequence for positive clone pBT40-4.4 is given in
CC (see R66204 FT).
SQ Sequence 44 BP; 12 A; 7 C; 12 G; 13 T;

Query Match 58.0%; Score 11.6; DB 1; Length 44;
Best Local Similarity 64.7%; Pred. No. 7e+02;

Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 atnccrtangcngtrt 17
|||:|||||:
DB 2 ATCCTGTAGCTGTGT 18

RESULT 13

ID Q79195 standard; DNA; 48 BP.

AC Q79195;

DT 04-AUG-1995 (first entry)

DE Oligo probe TAP113s for bovine tracheal antimicrobial peptide gene.

KW Antimicrobial peptide; TAP; probe; ss.

OS Synthetic.

PN M09426106-A.

PD 24-NOV-1994.

PF 11-MAY-1994; U05257.

PR 11-MAY-1993; US-060822.

PA (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.

PI Bevins CL, Diamond G, Zasloff MA;

DR WPI; 95-006233/01.

PT New antimicrobial peptide and precursor from mammalian trachea -

PT and related DNA, vector and transformed cells, useful as

PT pharmaceuticals and disinfectants

PS Example; page 27; 69 pp; English.

CC A mammalian TAP (R66203) was isolated by a combination of

CC size-exclusion, ion-exchange, and reverse-phase chromatographic

CC fractionations using antimicrobial activity against a strain of

CC E. coli as a functional assay. The mol. wt. of TAP is 4085 Da. The

CC predicted pI is 13.0 and there are no aromatic residues. A cDNA

CC sequence corresp. to the precursor peptide was cloned (see Q79533/
CC R66204), and the genomic TAP gene (Q79534/R66205) was isolated.

CC PCR primers for amplifying the TAP gene from bovine genomic DNA
CC were TAP69s and TAP610a. Oligo probes for hybridisation were

CC TAP48a, TAP54s, TAP160a, TAP113s, TAP118s.
SQ Sequence 48 BP; 13 A; 8 C; 14 G; 13 T;

Query Match 58.0%; Score 11.6; DB 1; Length 48;

Best Local Similarity 64.7%; Pred. No. 7.1e+02;

Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 atnccrtangcngtrt 17
|||:|||||:
DB 2 ATCCTGTAGCTGTGT 18

RESULT 14

ID T74158 standard; DNA; 29 BP.

AC T74158;

DT 29-SEP-1997 (first entry)

DE Salmonella enteritidis tct hybridisation oligonucleotide MKS3.

KW Enteropathogenic bacteria; enterobacteria; S.enteritidis; antibody;

OS Synthetic.

PN US5635617-A.

PD 03-JUN-1997.

PF 26-APR-1993; 054452.

PR 26-APR-1993; US-054452.

PA (UVVI-) UNIV VICTORIA INNOVATION & DEV CORP.

PI Clouthier SC, Collinson SK, Doran JL, Kay MW;

DR WPI; 97-309886/28.

PT Isolated Salmonella gene agfa - used for diagnosis of Salmonella or

PT enteropathogenic bacteria of the Enterobacteria family

PS Example 6; Column 37; 85pp; English.

CC The present sequence represents an oligonucleotide MKS3 that hybridises

CC to WK-10, which is used as part of the method in sequencing the tctCBA

CC gene cluster of Salmonella enteritidis. The nucleic acids produced can

CC be used to provide diagnostic assays for Salmonella and/or

CC enteropathogenic bacteria of the family Enterobacteria. They can also

CC be used to provide proteins and antibodies which can be used for assays.

CC The nucleic acid sequences can be used to provide probes or primers

CC which can specifically hybridise to nucleic acid molecules from greater

CC than 99% of Salmonella strains that are pathogenic to warm-blooded

CC animals relative to nucleic acid molecules from virtually all other

CC microbial organisms.
SQ Sequence 29 BP; 7 A; 4 C; 2 G; 16 T;

Query Match 57.0%; Score 11.4; DB 1; Length 29;

Best Local Similarity 60.0%; Pred. No. 8.6e+02;

Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 atnccrtangcngtrt 20
|||:|||||:
DB 8 ATTCACAAAGCTTTT 27

RESULT 15

ID V22156 standard; cDNA; 34 BP.

AC V22156;

DT 20-JUL-1998 (first entry)

DE BH3 Interacting domain death agonist polynucleotide fragment 17.

KW Human; BH3 Interacting domain death agonist; BID; BCL-2 family;

KW Apoptosis; regulation; cell death; inflammation; cancer; arthritis;

OS autoimmune disease; viral infection; lymphoproliferative; ss.

PN M09809980-A1.

PD 12-MAR-1998.

PF 09-SEP-1997; U15872.

PR 09-SEP-1996; US-706741.

PA (UNIM) UNIV WASHINGTON.

PI Korsmeyer SJ;

DR WPI; 98-193546/17.

PT BH3 interacting domain death agonist polypeptide - used for treating
PT decreased apoptotic conditions resulting from inflammation etc.
PS Disclosure; Page 23; 118pp; English.
CC The present sequence represents a BH3 interacting domain death agonist
CC (BID) polynucleotide fragment given in the present invention. The
CC protein, the DNA encoding it or antisense sequences can be used for
CC preventing or treating a decreased apoptotic state of a cell. The
CC decreased apoptotic state that is treated results from a disease such as
CC cancer, viral infections, lymphoproliferative conditions, arthritis,
CC inflammation and autoimmune diseases. Antibodies against the BID protein
CC can be used for detecting a BID polypeptide in a cell or population of
CC cell. The nucleic acid sequence and the BID protein can also be used for
CC treating immunodeficiency disease (including AIDS), senescence,
CC neurodegenerative disease, ischemic and reperfusion cell death,
CC infertility and wound-healing. Primers derived from the nucleic acid
CC encoding the BID protein can be used for detecting/quantitating the
CC protein and for detecting alterations in the nucleic acid encoding the
CC BID protein.
SQ Sequence 34 BP; 12 A; 9 C; 7 G; 6 T;

Query Match

57.0%; Score 11.4; DB 1; Length 34;

Best Local Similarity 60.0%; Pred. No. 8, 8e+02;

Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 atnccrtangcngtrtttctt 20

DB 22 ATWCCGATGATGATCTTCTT 3

Search completed: June 4, 2000, 16:24:04
Job time: 28900 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:02 ; Search time 4521.53 Seconds
(without alignments)
17.929 Million cell updates/sec

Title: US-09-164-714-5

Perfect score: 20
Sequence: 1 atncrtangcngtcttlytt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST: *
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: em_est20:*
21: em_est21:*
22: em_est22:*
23: em_est23:*
24: em_est24:*
25: em_est25:*
26: em_est26:*
27: em_est27:*
28: em_est28:*
29: em_est29:*
30: em_est30:*
31: em_est31:*
32: em_est32:*
33: em_est33:*
34: em_est34:*
35: em_est35:*
36: em_est36:*
37: em_est37:*
38: em_est38:*
39: em_est39:*
40: em_est40:*
41: em_est41:*
42: em_est42:*
43: em_est43:*
44: em_est44:*
45: em_est45:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*
106: gb_est87:*
107: gb_est88:*
108: gb_est89:*
109: gb_est90:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result

Query

No.	Score	Match	Length	DB	ID	Description
1	12.6	63.0	88	45	AI339161	AI339161 qtl4h01.x
2	12	59.0	72	69	AW116071	AW116071 f106g06.x
3	11.8	59.0	67	23	H41775	H41775 yn95c08.s1
4	11.8	59.0	79	28	AA116248	AA116248 mql1g10.r
5	11.6	58.0	52	51	AT749415	AT749415 at23g11.x
6	11.6	58.0	74	21	T99376	T99376 y64f04.s1
7	11.6	58.0	93	62	AI929034	AI929034 au64c11.y
8	11.6	58.0	94	84	B35774	B35774 HS-1030-BI-
9	11.6	58.0	95	33	AA449805	AA449805 zx09h01.s
10	11.4	57.0	58	20	D25841	D25841 HMG504213
11	11.4	57.0	80	23	H55643	H55643 CHR220582.C
12	11.4	57.0	84	44	AI318353	AI318353 ta76g04.x
13	11.4	57.0	87	44	AI318677	AI318677 ta49e05.x
14	11.4	57.0	100	40	AA910722	AA910722 ck84e12.s
15	11.4	57.0	100	42	AI109871	AI109871 vg80b10.r
16	11.2	56.0	40	36	AA601989	AA601989 nc08h07.s
17	11.2	56.0	81	79	AW250425	AW250425 2822308.3
18	11.2	56.0	89	45	AI385053	AI385053 fb06e09.x
19	11.2	56.0	91	42	AI130098	AI130098 SMOVL3CAN
20	11.2	56.0	100	37	AA706875	AA706875 z121f09.s
21	11.2	55.0	51	49	AI633437	AI633437 to46h07.x
22	11	55.0	52	49	AT661448	AT661448 va28d11.x
23	11	55.0	64	24	H80115	H80115 yu09g05.s1
24	11	55.0	64	63	AI950653	AI950653 wx51h01.x
25	11	55.0	67	62	AI921937	AI921937 wn86h08.x
26	11	55.0	68	29	AA153886	AA153886 mq50e02.r
27	11	55.0	76	27	AA052289	AA052289 mb42e11.r
28	11	55.0	78	33	AA412487	AA412487 zt99a10.s
29	11	55.0	84	46	AI463806	AI463806 va29d11.y
30	11	55.0	85	28	C21053	C21053 HDMG5000255
31	11	55.0	87	50	AI709401	AI709401 as37c10.x
32	11	55.0	88	44	AI318471	AI318471 ta43c04.x
33	11	55.0	88	63	AI956631	AI956631 ul79f04.y
34	11	55.0	89	44	AI318524	AI318524 ta79h04.x
35	11	55.0	90	43	AI204814	AI204814 zf-est134
36	11	55.0	90	45	AI357740	AI357740 qv73c05.x
37	11	55.0	94	44	AI282643	AI282643 qt70a06.x
38	11	55.0	95	29	AA155036	AA155036 ms14e01.r
39	11	55.0	96	46	AI442889	AI442889 sa28c02.x
40	11	55.0	98	37	AA710541	AA710541 vt50h03.r
41	10.8	54.0	43	41	AU006943	AU006943 AU006943
42	10.8	54.0	43	41	AU010101	AU010101 AU010101
43	10.8	54.0	43	41	AU010103	AU010103 AU010103
44	10.8	54.0	58	38	AA767515	AA767515 oa44a03.s
45	10.6	53.0	31	43	AI193015	AI193015 qe69a08.x

ALIGNMENTS

```

RESULT 1
LOCUS AI339161 88 bp mRNA EST 29-DEC-1998
DEFINITION qtl4h01.x1 NCI-CGAP GC4 Homo sapiens cDNA clone IMAGE:1947601 3'
SUBMIT 1; mRNA sequence.
ACCESSION AI339161
VERSION AI339161.1 GI:4076088
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 88)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_L_Strausberg@nih.gov

```

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLNt at: www.bio.lnlnl.gov/db/nci/image/image.html

FEATURES

Trace considered overall poor quality
 Seq primer: -40up from gibco
 High quality sequence stop: 1.
 Location/Qualifiers
 1..88
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1947601"
 /clone_1lb="NCI-CGAP-GC4"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker: 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
 19 a 16 c 21 g 32 t

Query Match
 Best local Similarity 63.0%; Score 12.6; DB 45; Length 88;
 Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 atnccrtangcgttrtytt 20
 ||||| |||||
 Db 18 ATCCCGTAGCAGGTTCAT 37

```

RESULT 2
LOCUS AW116071 72 bp mRNA EST 20-OCT-1999
DEFINITION f106g06.x1 Sugano Kawakami zebrafish DNA Danio rerio cDNA clone
ACCESSION AW116071
VERSION AW116071.1 GI:6082409
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
Cyprinidae; Cyprinidae; Rasbora; Danio.
REFERENCE 1 (bases 1 to 72)
AUTHORS Sugano,S., Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S.,
Hillier,L., Clifton,S., Allen,M., Gibbons,M., Jost,S., Kucaba,T.,
Martin,J., Pape,D., Steptoe,M., Underwood,K., Thelning,B.,
Ritter,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU Zebrafish EST Project 1999
JOURNAL Unpublished (1999)
COMMENT On Jul 7, 1999 this sequence version replaced g1:5866176.
Contact: S.L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
Sequencing by: Washington University Genome Sequencing Center
Seq primer: T7 ET from Amersham
High quality sequence stop: 55.
Location/Qualifiers

```


BASE COUNT 22 a 13 c 17 g 20 t 2 others

Query Match	58.0%	Score 11.6	DB 21	Length 74
Best Local Similarity	61.1%	Pred. No. 2e+04		
Matches 11, Conservative	2	Mismatches 5	Indels 0	Gaps 0

```

Qy      1 acnccrcangcngtrtcy 18
        ||  |  |  |  |  |  |
Db      3 ATTACATTAAGCTGTNTTC 20

```

RESULT	7
A1929034/c	
LOCUS	A1929034
DEFINITION	aue4c11.1 Schneider fetal brain mRNA
IMAGE:	2519540 5' , mRNA sequence.
ACCESSION	A1929034
VERSION	A1929034.1
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	1 (bases 1 to 93)
AUTHORS	Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Mattlin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE	WashU-NCI human EST project
JOURNAL	Unpublished (1997)
COMMENT	On Jun 5, 1998 this sequence version replaced gi:3188824.

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel.: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40RP from Gibco.

FEATURES	Location/Qualifiers
source	1. .93

5'-GAGAGAGAGAGAGAGACTCCAGAGATCCTTAATTAATTAATATATCCCCCCCCCCCC-3'
5' adaptor sequence:
5'-GAGAGAGAGAGAGACTCCAGTCTTTTCTTTTCTTTT-3'. The library was
size-selected for >0.5 Kb inserts and has an average
insert size estimated at 1.2 Kb. This library was
constructed using the CAP-trapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LNCIB-Area
Science Park, Trieste, Italy). "

ORIGIN

Query Match	58.08%	Score 11.6	DB 62,	Length 93,
Best Local Similarity	64.78%	Pred. No. 2e+04,		
Matches 11, Conservative	2,	Mismatches 4,	Indels 0,	Gaps 0,

Qy	1	atnccrtaengcgrttt	17
Db	49	atcccataagcttgct	33

LOCUS	B35774	94 bp	DNA	GSS	17-OCT-1997
DEFINITION	HS-1030-B1-E09-RR.abi C17 Human Genomic Sperm Library C Homo sapiens genomic clone Plate-CT810 Col-17 Row-J, genomic survey sequence.				

ACCESSION	B35774
VERSION	B35774.1
KEYWORDS	GI:2535143
SOURCE	GSS.
ORGANISM	human.
	Homo sapiens

REFERENCE
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 94)

AUTHORS
Malairares, G.G., Zackstone, K.D., Smith, T., Tipton, S., Schmidt, S.,
Tratcoff, R., Abajian, C., Blanchard, A., West, A., and Hood, L.E.

TITLE
Construction of a Characterized Clone Resource for Genomic
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors

JOURNAL
COMMENT
Unpublished (1997)
Contact: Mahiras GG, Zackrone KD, Hood L
University of Washington
Seattle, WA 98195, USA
Tel: (206) 616-8744
Fax: (206) 685-7301
Email: kzackrone@u.washington.edu
Sequence Tagged Connector
Plate: C7810 row: J column: 17
Class: BAC ends
High quality sequence stop: 94.

FEATURES	Location/Qualifiers
source	1. .94

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-Cr810 Col=17 Row=J"
/clone_1lb="CtR Human Genomic Sperm Library C"
/sex="M"
/note="Organ: sperm; Vector: pBeloBAC1; BAC clones in
E-coli DH10B"
BASE COUNT      18 a      20 c      13 g      41 t      2 others
ORIGIN

```

Query Match	58.0%;	Score 11.6;	DB 84;	Length 94;
Best Local Similarity	57.9%;	Pred. No. 2e+04;		
Matches 11; Conservative	3;	Mismatches	0;	Gaps 0;

```
QY      1 atncrtangcngtrtlyt 19
        |||:| | |||:| | |
Db      19 ATACTGTATCCTGTGTTT 37
```

[illegible]

Db 56 AGGCGTAGGCTGCTTT 75

RESULT 12
LOCUS AI318353
DEFINITION ta78604.x1 NCI-CGAP_HSC2 Homo sapiens CDNA clone IMAGE:2050230 3'
ACCESSION AI318353
VERSION AI318353.1 GI:4034235
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 84)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Herbert Morse, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

FEATURES
source
1. 84
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2050230"
/clone_1lb="NCI-CGAP_HSC2"
/tissue_type="stem cell 34+/38+"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: bone marrow; Vector: pAMP1; mRNA made from
bone marrow, stem cells 34+/38+, CDNA made by oligo-dT
priming. Directionally cloned. Size-selected on agarose
gel, average insert size 400 bp. Primary library,
non-amplified."

BASE COUNT 15 a 27 c 14 g 28 t
ORIGIN

Query Match 57.0%; Score 11.4; DB 44; Length 84;
Best Local Similarity 60.0%; Pred. NO. 2.6e+04;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 atnccrtangcngtrtytt 20
|||:|||||:|:|
Db 45 ATCCGATACGCGCTTCTT 64

RESULT 13
LOCUS AI318677
DEFINITION ta49905.x1 NCI-CGAP_Lu25 Homo sapiens CDNA clone IMAGE:2047424 3'
ACCESSION AI318677
VERSION AI318677.1 GI:4034612
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 87)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: I.M.A.G.E. Consortium, LLNL
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

FEATURES
source
1. 87
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2047424"
/clone_1lb="NCI-CGAP_Lu25"
/tissue_type="bronchioalveolar carcinoma"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: lung; Vector: pAMP1; mRNA made from lung
carcinoma tissue, CDNA made by oligo-dT priming.
Directionally cloned. Size-selected on agarose gel,
average insert size 500 bp. Primary library,
non-amplified."

BASE COUNT 21 a 29 c 14 g 23 t
ORIGIN

Query Match 57.0%; Score 11.4; DB 44; Length 87;
Best Local Similarity 60.0%; Pred. NO. 2.6e+04;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 atnccrtangcngtrtytt 20
|||:|||||:|:|
Db 48 ATCCATACGCGCTTCTT 67

RESULT 14
LOCUS AA910722
DEFINITION ok4e12.s1 NCI-CGAP_KiD3 Homo sapiens CDNA clone IMAGE:1520686 3'
ACCESSION AA910722
VERSION AA910722.1 GI:3050012
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 100)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
On Jan 19, 1998 this sequence version replaced gi:2152941.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbrrp/image/image.html

Trace considered overall poor quality
Insert Length: 370 Std Error: 0.00
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES

source
location/Qualifiers

1..100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1520686"
/clone_1ib="NCI CGAP_K1d3"
/lab_host="DH10B"
/note="organ: kidney; Vector: p7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dt) primer,
(double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified p7T3 vector. mRNA
source: 2 pooled kidneys. Library went through one round
of normalization. Library constructed by Bento Soares and
M. Fatima Bonaldo."

BASE COUNT 11 a 29 c 22 g 38 t

ORIGIN

Query Match 57.0%; Score 11.4; DB 40; Length 100;
Best local Similarity 60.0%; Pred. No. 2.6e+04;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 atnccrtangcgttrtytt 20

1 |||:| || |||:| |

Db 79 ACCCCATGAGCTGCTTCTT 98

RESULT 15

LOCUS A1098711

DEFINITION A1098711 100 bp mRNA EST 20-AUG-1998
vg80b10.r1 Barstead MPLRB1 Mus musculus CDNA clone IMAGE:872251 5',
similar to SW:COX1_RAT P05503 CYTOCHROME C OXIDASE POLYPEPTIDE I ;,
mRNA sequence.

ACCESSION A1098711

VERSION A1098711.1 GI:3448236

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 100)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT On Apr 14, 1993 this sequence version replaced gi:639245.
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:511731

Possible reversed clone: similarity on wrong strand

Seq primer: -28m3 rev2 ET from Amersham
High quality sequence stop: 1.

FEATURES

source
location/Qualifiers

1..100
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="IMAGE:872251"
/clone_1ib="Barstead MPLRB1"
/sex="mixed"
/tissue_type="kidney"
/dev_stage="6 weeks"
/lab_host="DH10B"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dt) primer [5',
TGTTCGATCTGAGTGGAGCGCCGCCCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[CATGCAATTCGTACC], digested with Not I and cloned into the
Not I and Eco RI sites of the modified p7T3 vector.
Library constructed by Bob Barstead."

BASE COUNT 33 a 22 c 11 g 34 t

ORIGIN

Query Match 57.0%; Score 11.4; DB 42; Length 100;
Best local Similarity 60.0%; Pred. No. 2.6e+04;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 atnccrtangcgttrtytt 20

1 |||:| || |||:| |

Db 9 ATTTCGTACGCTCTATATGTT 28

Search completed: June 4, 2000, 13:53:05
Job time: 20929 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:39 ; Search time 244.64 Seconds
(Without alignments)
10.627 Million cell updates/sec

Title: US-09-164-714-5

Perfect score: 20

Sequence: 1 atnccrtangcngtrtlytt 20

Scoring table: IDENTITY_NUC

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/6.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/PCBUS.COMB.seq:*
7: /cgn2_6/ptodata/2/1na/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	12.6	63.0	97	3	US-08-852-807-17 Sequence 17, Appl
C 2	12.2	61.0	48	1	US-08-060-822A-18 Sequence 18, Appl
C 3	12.2	61.0	48	6	PCT-US94-05257-18 Sequence 18, Appl
4	12.2	61.0	77	1	US-08-399-412A-18 Sequence 2, Appl
5	12	60.0	49	3	US-07-673-661B-2 Sequence 50, Appl
C 6	11.8	59.0	61	4	US-08-483-528B-50 Sequence 50, Appl
C 7	11.8	59.0	61	5	US-08-673-799C-50 Sequence 135, App
C 8	11.8	59.0	87	1	US-08-425-336-135 Sequence 135, App
C 9	11.8	59.0	87	2	US-08-488-113B-135 Sequence 135, App
C 10	11.8	59.0	87	2	US-08-477-484B-135 Sequence 135, App
C 11	11.8	59.0	87	2	US-08-107-669D-60 Sequence 60, Appl
C 12	11.8	59.0	87	2	US-08-472-788A-60 Sequence 60, Appl
C 13	11.8	59.0	87	2	US-08-477-531B-60 Sequence 60, Appl
C 14	11.8	59.0	87	3	US-08-646-360-135 Sequence 135, App
C 15	11.8	59.0	87	3	US-08-082-842A-60 Sequence 60, Appl
C 16	11.6	58.0	44	1	US-07-603-451A-10 Sequence 10, Appl
C 17	11.6	58.0	44	1	US-08-060-822A-11 Sequence 11, Appl
C 18	11.6	58.0	44	6	PCT-US94-05257-11 Sequence 11, Appl
C 19	11.6	58.0	48	1	US-08-060-822A-19 Sequence 19, Appl
C 20	11.6	58.0	48	6	PCT-US94-05257-19 Sequence 19, Appl
C 21	11.4	57.0	29	1	US-08-233-788A-8 Sequence 8, Appl
C 22	11.4	57.0	34	4	US-08-924-695A-17 Sequence 17, Appl
C 23	11.4	57.0	40	3	US-08-281-423-7 Sequence 7, Appl
C 24	11.2	56.0	39	3	US-08-470-419-6 Sequence 6, Appl
C 25	11.2	56.0	39	3	US-08-761-828-6 Sequence 6, Appl
C 26	11.2	56.0	39	4	US-08-290-105-6 Sequence 6, Appl
C 27	11.2	56.0	39	5	US-08-776-949-6 Sequence 6, Appl

ALIGNMENTS

28	11.2	56.0	48	3	US-08-470-419-20	Sequence 20, Appl
29	11.2	56.0	48	3	US-08-761-828-20	Sequence 20, Appl
30	11.2	56.0	48	4	US-08-290-105-20	Sequence 20, Appl
31	11.2	56.0	48	5	US-08-776-949-20	Sequence 378, App
32	11.2	56.0	62	6	PCT-US96-09455A-378	Sequence 97, Appl
33	11	55.0	27	1	US-08-399-412A-97	Sequence 107, App
C 34	11	55.0	29	2	US-08-623-891-107	Sequence 3, Appl
C 35	11	55.0	63	1	US-07-901-708-3	Sequence 11, App
C 36	10.6	53.0	22	3	US-08-117-952-111	Sequence 26, Appl
37	10.6	53.0	29	1	US-08-495-743-26	Sequence 26, Appl
38	10.6	53.0	29	2	US-08-495-739-26	Sequence 26, Appl
39	10.6	53.0	29	2	US-08-495-741-26	Sequence 5, Appl
40	10.6	53.0	30	3	US-08-583-276-5	Sequence 70, Appl
41	10.6	53.0	31	2	US-08-403-762A-70	Sequence 4, Appl
C 42	10.6	53.0	31	3	US-08-829-961-4	Sequence 8, Appl
C 43	10.6	53.0	34	5	US-08-886-466-8	Sequence 4, Appl
C 44	10.6	53.0	40	5	US-08-841-267-4	Sequence 14, Appl
45	10.6	53.0	77	1	US-08-399-412A-14	

RESULT 1
US-08-852-807-17/C
; Sequence 17, Application US/08852807
; Patent No. 5861298
; GENERAL INFORMATION:
; APPLICANT: Debouck, Christine
; APPLICANT: Drake, Fred
; APPLICANT: Gowen, Maxine
; APPLICANT: Hood, Julie
; APPLICANT: Hastings, Gregg
; APPLICANT: Adams, Mark
; APPLICANT: Fraser, Claire
; APPLICANT: Lee, No. 5861298man
; APPLICANT: Kirness, Even
; APPLICANT: Blake, Judith
; APPLICANT: Fitzgerald, Lisa
; TITLE OF INVENTION: CATHEPSIN K GENE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA: 807
; APPLICATION NUMBER: US/08/852, 807
; FILING DATE: 07-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019, 942
; FILING DATE: 14-JUNE-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/020, 273
; FILING DATE: 17-JUNE-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/026, 273
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50006-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090

TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
US-08-852-807-17

Query Match 63.0%; Score 12.6; DB 3; Length 97;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 atncrtangngtrtlytt 20
|||: ||| |||: |||
Db 69 ATCCACAGCGTGTCTT 50

RESULT 2
US-08-060-822A-18/c
; Sequence 18, Application US/08060822A
; Patent No. 5432270
; GENERAL INFORMATION:
; APPLICANT: Bevins, Charles L.
; APPLICANT: Diamond, Gill
; APPLICANT: Zaslloff, Michael
; TITLE OF INVENTION: No. 5432270e1 Antimicrobial Peptides
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5432270ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/060,822A
; FILING DATE: 19930511
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/991,200
; FILING DATE: December 15, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca Ralph Gaumond
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-060-822A-18

Query Match 61.0%; Score 12.2; DB 1; Length 48;
Best Local Similarity 61.1%; Pred. No. 2.1e+02;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 atncrtangngtrtlytt 18
|||: ||| |||: |||
Db 47 ATCCTGTAACTGTGT 30

RESULT 3
PCT-US94-05257-18/c
; Sequence 18, Application PC/TUS9405257
; GENERAL INFORMATION:
; APPLICANT: Bevins, Charles L.
; APPLICANT: Diamond, Gill
; APPLICANT: Zaslloff, Michael
; TITLE OF INVENTION: Novel Antimicrobial Peptides
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05257
; FILING DATE: 11-MAY-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/060,822
; FILING DATE: May 11, 1993
; APPLICATION NUMBER: 07/991,200
; FILING DATE: December 15, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca Ralph Gaumond
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US94-05257-18

Query Match 61.0%; Score 12.2; DB 6; Length 48;
Best Local Similarity 61.1%; Pred. No. 2.1e+02;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 atncrtangngtrtlytt 18
|||: ||| |||: |||
Db 47 ATCCTGTAACTGTGT 30

RESULT 4
US-08-399-412A-18
; Sequence 18, Application US/08399412A
; Patent No. 5622828
; GENERAL INFORMATION:
; APPLICANT: Patma, David
; APPLICANT: Gold, Larry
; TITLE OF INVENTION: High-Affinity Oligonucleotide
; TITLE OF INVENTION: Ligands To Secretory Phospholipase
; NUMBER OF SEQUENCES: 122
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.

STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,412A
FILING DATE: 6-MARCH-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: Julie L. Bernard
REGISTRATION NUMBER: 36,450
REFERENCE/DOCKET NUMBER: NEX27
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-399-412A-18

Query Match 61.0%; Score 12.2; DB 1; Length 77;
Best Local Similarity 27.8%; Pred. No. 2.3e+02;
Matches 5; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY 2 tncrtangcngtrtlyt 19
: ||:| | | : : : : :
Db 35 UACCGUAGACGUAVUCG 52

RESULT 5
US-07-673-661B-2
Sequence 2, Application US/07673661B
Patent No. 5843667
GENERAL INFORMATION:
APPLICANT: Weisburg, William G.
TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE DETECTION OF
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENE-TRAK SYSTEMS
STREET: 31 New York Avenue
CITY: Framingham
STATE: Massachusetts
COUNTRY: USA
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch/1.44MB Macintosh formatted diskette
COMPUTER: Macintosh Itcl
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft WORD (Version 5.1a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/673,661B
FILING DATE: MARCH 22, 1991

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: N/A
FILING DATE: N/A
ATTORNEY/AGENT INFORMATION:
NAME: Janluk, Anthony J.
REGISTRATION NUMBER: 29,809
REFERENCE/DOCKET NUMBER: F-120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-3113
TELEFAX: (508) 879-6462
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 NUCLEOTIDES
TYPE: NUCLEIC ACID SEQUENCE
STRANDEDNESS: SINGLE STRANDED
TOPOLOGY: LINEAR
MOLECULE TYPE: DNA PROBE
DESCRIPTION: THIS SEQUENCE IS USEFUL AS A DNA PROBE OR
DESCRIPTION: AS AN RNA PROBE FOR THE SPECIFIC DETECTION OF UREAPLASMA
FEATURE:
NAME/KEY: GTS OLIGO NUMBER 2256
US-07-673-661B-2

Query Match 60.0%; Score 12; DB 3; Length 49;
Best Local Similarity 70.6%; Pred. No. 2.8e+02;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 atncrtangcngtrtct 17
: ||:| | | | | : : : : :
Db 33 ATGCCCTATGCGGTAT 49

RESULT 6
US-08-483-528B-50/c
Sequence 50, Application US/08483528B
Patent No. 5939532
GENERAL INFORMATION:
APPLICANT: NAKAMURA, KAZUYASU
APPLICANT: KOIKE, MASAMICHI
APPLICANT: SHITARA, KENYA
APPLICANT: HANAI, NOBUO
APPLICANT: KUMANA, YOSHIRISA
APPLICANT: HASEGAWA, MAMORU
TITLE OF INVENTION: HUMANIZED ANTIBODIES
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYTE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,528B
FILING DATE: 07-JUN-95
CLASSIFICATION: 536
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4000
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 base pairs
TYPE: nucleic acid
STRANDEDNESS: double


```

;
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,113B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/425,336
; FILING DATE: 18-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-488-113B-135

Query Match          59.0%; Score 11.8; DB 2; Length 87;
Best Local Similarity 66.7%; Pred. No. 3.8e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      6 rtangcngtrtlytt 20
       :|||1111111111
Db      44 ATAGGCAAGTGTCTT 30

RESULT 10
US-08-477-484B-135/C
; Sequence 135, Application US/08477484B
; Patent No. 5756699
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnicka, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,484B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/425,336
; FILING DATE: 18-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-477-484B-135

Query Match          59.0%; Score 11.8; DB 2; Length 87;
Best Local Similarity 66.7%; Pred. No. 3.8e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      6 rtangcngtrtlytt 20
       :|||1111111111
Db      44 ATAGGCAAGTGTCTT 30

RESULT 11
US-08-107-669D-60/C
; Sequence 60, Application US/08107669D
; Patent No. 5766886
; GENERAL INFORMATION:
; APPLICANT: Studnicka, Gary M.
; TITLE OF INVENTION: Modified Antibody Variable Domains (as amended)
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/107,669D
; FILING DATE: 13-AUG-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
```

```

1 APPLICATION NUMBER: PCT/US92/10906
2
3 FILING DATE: 14-DEC-1992
4
5 PRIOR APPLICATION DATA:
6
7 APPLICATION NUMBER: US 07/808,464
8
9 FILING DATE: 13-DEC-1991
10
11 ATTORNEY/AGENT INFORMATION:
12
13 NAME: Michele A. Cimbalà
14
15 REGISTRATION NUMBER: 33,851
16
17 REFERENCE/DOCKET NUMBER: 0610.1000001/MAC
18
19 TELECOMMUNICATION INFORMATION:
20
21 TELEPHONE: 202/371-2600
22
23 TELEFAX: 202/371-2540
24
25 INFORMATION FOR SEQ ID NO: 60:
26
27 SEQUENCE CHARACTERISTICS:
28
29 LENGTH: 87 base pairs
30
31 TYPE: nucleic acid
32
33 STRANDEDNESS: single
34
35 TOPOLOGY: linear
36
37 MOLECULE TYPE: DNA
38
39 US-08-107-669D-60

```

Query Match	59.0%;	Score 11.8;	DB 2;	Length 87;
Best Local Similarity	66.7%;	Pred. No. 3.8e+02;		
Matches 10; Conservative	3;	Mismatches 2;	Indels 0;	Gaps 0

QY	6	rtangcngtrctt	20
		: :	
Db	44	ATAGGCAGTGTCTT	30

RESULT 12
 US-08-472-788A-60/C
 Sequence 60, Application US/08472788A
 Patent No. 5770196
 GENERAL INFORMATION:
 APPLICANT: Studnicka, Gary M.
 TITLE OF INVENTION: Modified Antibody Variable Domains
 NUMBER OF SEQUENCES: 89
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Stearns, Kessler, Goldstein and Fox P.L.L.C.
 STREET: 1100 New York Ave., N.W., Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/472,788A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/082,842
 FILING DATE: 23-JUN-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92/10906
 FILING DATE: 14-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/808,464
 FILING DATE: 13-DEC-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Cimdala, Michele A.
 REGISTRATION NUMBER: 33,851
 REFERENCE/DOCKET NUMBER: 0610.1000003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202/371-2600
 TELEFAX: 202/371-2540
 TELEX:
 INFORMATION FOR SEQ ID NO: 60:

```

; SEQUENCE CHARACTERISTICS
; LENGTH: 87 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-472-788A-60

```

Query Match	59.0%;	Score 11.8;	DB 2;	Length 87;
Best Local Similarity	66.7%;	Pred. No. 3.8e+02;		
Matches 10;	Conservative 3;	Mismatches 2;	Indels 0;	Gaps 0;

QY	6	rtangcngtrtlylt	20
		: : :	
Db	44	ATAGCAGTGTCTT	30

```

RESULT 13
US-08-477-531B-60/c
: Sequence 60, Application US/08477531B
: Patent No. 5821123
: GENERAL INFORMATION:
: APPLICANT: Studnicka, Gary M.
: TITLE OF INVENTION: Modified Antibody Variable Domains (as amended)
: NUMBER OF SEQUENCES: 67
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.
: STREET: 1100 New York Ave., N.W., Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: United States of America
: ZIP: 20005-3934
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/477,531B
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 436
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/107,669
: FILING DATE: 13-AUG-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/10906
: FILING DATE: 14-DEC-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/808,464
: FILING DATE: 13-DEC-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Michele A. Cimbalia
: REGISTRATION NUMBER: 33,851
: REFERENCE/DOCKET NUMBER: 0610.1000004/MAC
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202/371-2600
: TELEFAX: 202/371-2540
: INFORMATION FOR SEQ. ID NO: 60:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 87 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: US-08-477-531B-60

```

Query Match	59.0%	Score 11.8;	DB 2;	Length 87;
Best Local Similarity	66.7%;	Pred. No. 3.8e+02;		
Matches 10; Conservative	3;	Mismatches 2;	Indels 0;	Gaps 0;
0y 6 rtangcngtrttt	20			

```

Db      44  ATAGCAGTGTCTT  30
          :||| ||: ||: ||

```

RESULT 14
 US-08-646-360-135/c
 ; Sequence 135 Application US/08646360
 ; Patent No. 5837491
 GENERAL INFORMATION:
 APPLICANT: Better, Marc D.
 APPLICANT: Carroll, Stephen F.
 APPLICANT: Studnika, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 TITLE OF INVENTION: Proteins
 NUMBER OF SEQUENCES: 173
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/646,360
 FILING DATE: 13-MAY-1996
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/05348
 FILING DATE: 12-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/064,691
 FILING DATE: 12-MAY-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/988,430
 FILING DATE: 09-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McNicholas, Janet M.
 REGISTRATION NUMBER: 32,918
 REFERENCE/DOCKET NUMBER: 200-70.P4
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/707-8889
 TELEFAX: 312/707-9155
 TELEX: 650 388-1248
 INFORMATION FOR SEQ ID NO: 135:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 87 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 US-08-646-360-135

Query Match	59.0%	Score 11.8;	DB 3	Length 87;
Best Local Similarity	66.7%	Pred. NO. 3	8e02	
Matches	10;	Conservative	3;	Mismatches 2;
				Indels 0;
				Gaps 0.
OY	6	rtangcngtrttt	ttc	20
		..		
Db	44	ATVAGCAGTGTCTT	CTT	30

```

15
US-08-082-842A-60/c
: Sequence 60, Application US/08082842A
: Patent No. 3869619
: GENERAL INFORMATION:
: APPLICANT: Studnicka, Gary M.
: TITLE OF INVENTION: Modified Antibody Variable Domains
: NUMBER OF SEQUENCES: 89
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.
: STREET: 1100 New York Ave., N.W., Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3934
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/082,842A
: FILING DATE: 23-JUN-1993
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/10906
: FILING DATE: 14-DEC-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/808,464
: FILING DATE: 13-DEC-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Cimbal, Michele A.
: REGISTRATION NUMBER: 33,851
: REFERENCE/DOCKET NUMBER: 0610.1000002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202/371-2600
: TELEFAX: 202/371-2540
: TELEX:
: INFORMATION FOR SEQ ID NO: 60:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 87 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: US-08-082-842A-60

```

Query Match	59.0%	Score 11.8	DB 3	Length 87
Best Local Similarity	66.7%	Pred. No. 3,8e+02		
Matches 10, Conservative	3	Mismatches 2	Indels 0	Gaps 0
OY	6	rtangcngtrtttyt	20	
			
Db	44	ATAGGCACTGTCTT	30	

Search completed: June 4, 2000, 16:09:40
Job time: 28059 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:30 ; Search time 1236.38 Seconds
(without alignments)
-22.031 Million cell updates/sec

Title: US-09-164-714-8
Perfect score: 28
Sequence: 1 ccctatgtgtgtcccaaatgtgtcaag 28

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_bai: *
2: gb_baz: *
3: gb_om: *
4: gb_ov: *
5: gb_pat: *
6: gb_ph: *
7: gb_pil: *
8: gb_pil2: *
9: gb_prl: *
10: gb_prl2: *
11: gb_prl3: *
12: gb_ro: *
13: gb_sts: *
14: gb_sy: *
15: gb_un: *
16: gb_vl: *
17: em_fun: *
18: em_hum1: *
19: em_hum2: *
20: em_in: *
21: em_om: *
22: em_or: *
23: em_ov: *
24: em_pat: *
25: em_ph: *
26: em_pl: *
27: em_ro: *
28: em_sts: *
29: em_sy: *
30: em_un: *
31: em_vl: *
32: gb_htg1: *
33: gb_htg2: *
34: gb_in1: *
35: gb_in2: *
36: em_bai: *
37: em_baz: *
38: em_hum3: *
39: em_hum4: *
40: gb_prl4: *
41: gb_htg3: *
42: gb_htg4: *
43: gb_htg5: *
44: gb_htg6: *

45: gb_htg7: *
46: em_htg1: *
47: em_htg2: *
48: em_htg3: *
49: em_hum5: *
50: gb_pil3: *
51: gb_prl5: *
52: gb_htg8: *
53: gb_htg9: *
54: gb_htg10: *
55: gb_htg11: *
56: gb_htg12: *
57: gb_htg13: *
58: gb_htg14: *

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	16	57.1	39	5	AR012248	AR012248 Sequence
C 2	16	57.1	39	5	I14946	I14946 Sequence 32
C 3	16	57.1	39	5	I73666	I73666 Sequence 32
4	15	53.6	34	5	A39689	A39689 Sequence 2
5	15	53.6	34	5	A48961	A48961 Sequence 8
6	14.8	52.9	75	40	AF064903	AF064903 Homo sapi
7	14.8	52.9	92	10	HUMPHOS15	MS9733 Human muscl
8	14.2	50.7	53	4	FR08001	AU008001 Fugu rudr
9	14.2	50.7	60	8	AF101150	AF101150 Brassica
C 10	14.2	50.7	72	5	AR040932	AR040932 Sequence
C 11	14.2	50.7	72	5	AR051839	AR051839 Sequence
C 12	14.2	50.7	72	5	I36229	I36229 Sequence 31
C 13	14.2	50.7	90	9	GORINSORFA	M69096 Gorilla end
C 14	14.2	50.7	99	34	DDIDKE	MS9748 D.discoiden
15	14	50.0	27	5	E14994	E14994 PCR primer
16	14	50.0	36	5	A78745	A78745 Sequence 6
17	14	50.0	36	5	AR014718	AR014718 Sequence
18	13.8	49.3	35	5	AR055776	AR055776 Sequence
C 19	13.8	49.3	35	5	AR055777	AR055777 Sequence
20	13.8	49.3	51	12	U92178	U92178 Mus musculu
21	13.6	48.6	26	5	AR014478	AR014478 Sequence
C 22	13.6	48.6	59	10	S59972	S59972 TAL1d1-prot
C 23	13.6	48.6	60	10	S72312	S72312 phosphofruc
24	13.6	48.6	81	3	PIGPROG	M25036 Hog prognst
25	13.6	48.6	99	12	MUSLRP03	L13679 Mouse leuko
26	13.4	47.9	44	5	AR048165	AR048165 Sequence
27	13.4	47.9	77	13	RICL254A	D25437 Rice genom
C 28	13.4	47.9	87	10	S71607	S71607 pol clone
C 29	13.2	47.1	23	5	AR063157	AR063157 Sequence
C 30	13.2	47.1	29	5	AR003150	AR003150 Sequence
C 31	13.2	47.1	29	5	AR003243	AR003243 Sequence
C 32	13.2	47.1	29	5	AR009108	AR009108 Sequence
C 33	13.2	47.1	29	5	AR011395	AR011395 Sequence
C 34	13.2	47.1	29	5	AR028081	AR028081 Sequence
C 35	13.2	47.1	29	5	AR052797	AR052797 Sequence
C 36	13.2	47.1	29	5	I18033	I18033 Sequence 26
37	13.2	47.1	43	5	A05133	A05133 Oligonucleo
38	13.2	47.1	50	9	HEPHWIN2	K01660 Hepatitis B
C 39	13.2	47.1	72	5	I00803	I00803 Sequence 7
C 40	13.2	47.1	77	5	A00281	A00281 Modified fr
41	13.2	47.1	77	5	A16012	A16012 HSA oligonu
C 42	13.2	47.1	93	12	MMU66712	U86712 Mus musculu
C 43	13	46.4	21	5	AR024394	AR024394 Sequence
C 44	13	46.4	21	5	I71186	I71186 Sequence 38
C 45	13	46.4	21	5	I83745	I83745 Sequence 38

ALIGNMENTS


```
source 1..34
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 5 a 6 c 14 g 9 t
ORIGIN

Query Match 53.6%; Score 15; DB 5; Length 34;
Best Local Similarity 78.3%; Pred. No. 7.2e+03;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ccatgtgtgtgccaaattgtg 23
||| ||||| ||| |||||
Db 7 CCTTACTGTGGTGCACGATTGG 29

RESULT 6
AF064903 75 bp mRNA PRI 03-MAR-1999
LOCUS Homo sapiens clone 12-1 immunoglobulin heavy chain alpha VDJ region
DEFINITION mRNA, partial cds.
ACCESSION AF064903
VERSION AF064903.1 GI:4321896
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 75)
AUTHORS Rowley,A.H., Eckerley,C.A., Shulman,S.T. and Baker,S.C.
TITLE Evidence of an antigen-driven Iga immune response in the vascular
wall in acute Kawasaki Syndrome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 75)
AUTHORS Rowley,A.H., Eckerley,C.A., Shulman,S.T. and Baker,S.C.
TITLE Direct Submission
JOURNAL Submitted (14-MAR-1998) Pediatrics, Loyola University, 2160 S First
Avenue, Maywood, IL 60153, USA
FEATURES
source
1..75
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="12-1"
/tissue="vascular"
/note="from aorto-iliac junction of a 10-year-old child
with fatal acute Kawasaki Syndrome"
<1..>75
/note="includes IGHV3 gene"
/codon_start=1
/product="immunoglobulin heavy chain alpha VDJ region"
/protein_id="AADI5867.1"
/db_xref="GI:4321897"
/translation="AVYYCAKNQOVTPPYHYVMDVWG"

BASE COUNT 19 a 22 c 15 g 19 t
ORIGIN

Query Match 52.9%; Score 14.8; DB 40; Length 75;
Best Local Similarity 73.1%; Pred. No. 9.4e+03;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 ccatgtgtgtccaaattgtgcaag 28
||||| ||||| ||||| |||||
Db 6 CTATTACTGTGCGAATAATGCTCAAG 31

RESULT 7
HUMPHOSIS 92 bp DNA PRI 05-DEC-1995
LOCUS Human muscle-type phosphofructokinase (PFK-M) gene, exon 16.
DEFINITION M59733
ACCESSION M59733.1 GI:188647
VERSION
KEYWORDS muscle-type phosphofructokinase.
```

```
SEGMENT 15 of 23
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 92)
AUTHORS Yamasaki,T., Nakajima,H., Kono,N., Hotta,K., Yamada,K., Imai,E.,
Kuwaita,M., Noguchi,T., Tanaka,T. and Tani,S.
TITLE Structure of the entire human muscle phosphofructokinase-encoding
JOURNAL Gene 104 (2), 277-282 (1991)
MEDLINE 92009225
FEATURES
source
1..92
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1cen-q32"
16..86
/feature="PFK-M"
/note="G00-120-277"
/number=16

BASE COUNT 23 a 17 c 32 g 20 t
ORIGIN

Query Match 52.9%; Score 14.8; DB 10; Length 92;
Best Local Similarity 73.1%; Pred. No. 9.5e+03;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 ccatgtgtgtccaaattgtgcaag 28
||||||| ||| ||||| |||
Db 36 CTATGTTGGGGCTGGACTGGCCCAAG 61
```

```
RESULT 8
FRU8001/C
LOCUS FRU8001 53 bp DNA VRT 18-JAN-1999
DEFINITION Fugu rubripes Eya4 gene, exon 4.
ACCESSION AJ008001
VERSION AJ008001.1 GI:4164315
KEYWORDS EYA4 gene.
SOURCE Fugu rubripes.
ORGANISM Fugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
Tetraodontiformes; Tetraodontidae; Tetraodontidae; Fugu.
REFERENCE 1 (bases 1 to 53)
AUTHORS Hanson,I.M.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1998) Hanson I.M., Cell Genetics Section, MRC
Human Genetics Unit, Western General Hospital, Crewe Road,
Edinburgh, EH4 2XU, UK
2 (bases 1 to 53)
REFERENCE Borsani,G., deGrandi,A., Ballabio,A., Bulfone,A., Bernard,L.,
Banfi,S., Gattuso,C., Mariani,M., Dixon,M., Donna,D., Metcalfe,K.,
Winter,R., Robertson,M., Axton,R., Brown,A., van Heyningen,V. and
Hanson,I.
TITLE EYA4, a novel vertebrate gene related to drosophila eyes absent
JOURNAL Hum. Mol. Genet. 8 (1), 11-23 (1999)
MEDLINE 99105912
FEATURES
source
1..53
Location/Qualifiers
/organism="Fugu rubripes"
/db_xref="taxon:31033"
/clone="cosmid 032C09"
1..53
/feature="Eya4"
/feature="Eya4"
1..>53
/feature="Eya4"
/number=4
/usedin="AJ007998:Eya4_CDS"

BASE COUNT 14 a 11 c 12 g 16 t
ORIGIN
```

Query Match 50.7%; Score 14.2; DB 4; Length 53;
Best Local Similarity 84.2%; Pred. No. 1.8e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 atgttggtgccaattg 23
||||| ||||| ||
Db 49 ATGTGTGTGCGCAAAACGG 31

RESULT 9
AF101150 60 bp DNA PLN 29-MAR-1999
LOCUS Brassica napus clone na8 SINE S1 integration site sequence.
ACCESSION AF101150
VERSION AF101150.1 GI:4530252
KEYWORDS
SOURCE rape.
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; euroids II; Brassicales; Brassicaceae;
Brassica.
REFERENCE 1 (bases 1 to 60)
AUTHORS Tatout,C., Laveie,L. and Deragon,J.M.
TITLE Similar target site selection occurs in integration of plant and
mammalian retroposons
JOURNAL J. Mol. Evol. (1999) In press
REFERENCE 2 (bases 1 to 60)
AUTHORS Tatout,C., Laveie,L. and Deragon,J.M.
TITLE Direct Submision
JOURNAL Submitted (26-OCT-1998) UMR 6547 BIOMOVE, CNRS, Universite Blaise
Pascal, 24 Av. des landais, Aubliere 63177, France

FEATURES
source 1..60
/organism="Brassica napus"
/db_xref="taxon:3708"
/clone="na8"
misc_feature 28..35
/note="position of the direct repeats formed after S1
integration"

BASE COUNT 19 a 6 c 15 g 20 t
ORIGIN

Query Match 50.7%; Score 14.2; DB 8; Length 60;
Best Local Similarity 84.2%; Pred. No. 1.8e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 atgttggtgccaattg 23
||||| ||||| |||||
Db 31 ATGTGTGTGCGCAAAATTGC 49

RESULT 10
AR040932 72 bp DNA PAT 29-SEP-1999
LOCUS AR040932
DEFINITION Sequence 31 from patent US 5811238.
ACCESSION AR040932
VERSION AR040932.1 GI:5961428
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 72)
AUTHORS Stemmer,W.P.C. and Cramer,I.A.
TITLE Methods for generating polynucleotides having desired
characteristics by iterative selection and recombination
JOURNAL Patent: US 5811238-A 31 22-SEP-1999;
FEATURES location/Qualifiers
source 1..72
/organism="unknown"

BASE COUNT 21 a 17 c 18 g 16 t
ORIGIN

Query Match 50.7%; Score 14.2; DB 5; Length 72;
Best Local Similarity 70.4%; Pred. No. 1.8e+04;
Matches 19; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 ccatgttggtgccaattgcaag 28
||||| ||||| |||||
Db 65 CTTCGTGTGCGTACCACATATATGTAACG 39

RESULT 11
AR051839 72 bp DNA PAT 29-SEP-1999
LOCUS AR051839/c
DEFINITION Sequence 31 from patent US 5830721.
ACCESSION AR051839
VERSION AR051839.1 GI:5975203
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 72)
AUTHORS Stemmer,W.P.C. and Cramer,I.A.
TITLE DNA mutagenesis by random fragmentation and reassembly
JOURNAL Patent: US 5830721-A 31 03-NOV-1998;
FEATURES location/Qualifiers
source 1..72
/organism="unknown"

BASE COUNT 21 a 17 c 18 g 16 t
ORIGIN

Query Match 50.7%; Score 14.2; DB 5; Length 72;
Best Local Similarity 70.4%; Pred. No. 1.8e+04;
Matches 19; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 ccatgttggtgccaattgcaag 28
||||| ||||| |||||
Db 65 CTTCGTGTGCGTACCACATATATGTAACG 39

RESULT 12
I36229 72 bp DNA PAT 13-MAY-1997
LOCUS I36229/c
DEFINITION Sequence 31 from patent US 5605793.
ACCESSION I36229
VERSION I36229.1 GI:2086742
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 72)
AUTHORS Stemmer,W.P.C.
TITLE Methods for in vitro recombination
JOURNAL Patent: US 5605793-A 31 25-FEB-1997;
FEATURES location/Qualifiers
source 1..72
/organism="unknown"

BASE COUNT 21 a 17 c 18 g 16 t
ORIGIN

Query Match 50.7%; Score 14.2; DB 5; Length 72;
Best Local Similarity 70.4%; Pred. No. 1.8e+04;
Matches 19; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 ccatgttggtgccaattgcaag 28
||||| ||||| |||||
Db 65 CTTCGTGTGCGTACCACATATATGTAACG 39

RESULT 13
GORINSORF/c
LOCUS GORINSORF 90 bp DNA PRI 27-APR-1993
DEFINITION Gorilla endogenous retroviral DNA.
ACCESSION M69096
VERSION M69096.1 GI:177060
KEYWORDS
SOURCE
ORGANISM Gorilla DNA.
Gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.
REFERENCE
AUTHORS 1 (bases 1 to 90)
TITLE Shih,A., Coutavas,E.E. and Rush,M.G.
JOURNAL Evolutionary implications of primate endogenous retroviruses
MEDLINE Virology 182, 495-502 (1991)
91220699
FEATURES
source
1..90
/organism="Gorilla gorilla"
/db_xref="taxon:9593"
CDS
1..90
/partial
/note="ORF"
/codon_start=1
/protein_id="AA35469.1"
/db_xref="GI:177061"
/translation="EKNSPTLDEALHRDITDFTRRHPIITLLQ"
BASE COUNT 19 a 31 c 25 t
ORIGIN
Query Match 50.7%; Score 14.2; DB 9; Length 90;
Best Local Similarity 84.2%; Pred. No. 1.8e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 10 ggtgccaaatggtcaag 28
||||| ||||| |||||
Db 63 ggtgccaaatcactcang 45
RESULT 14
DDIDDK/c
LOCUS DDIDDK 99 bp DNA INV 27-APR-1993
DEFINITION D.discoidium protein kinase 5 gene, partial cds.
ACCESSION M59748
VERSION M59748.1 GI:167725
KEYWORDS protein kinase 5.
SOURCE Dictyostelium discoidium (strain AX-3) DNA.
ORGANISM Dictyostelium discoidium
Eukaryota; Dictyostelida; Dictyostelium.
REFERENCE
AUTHORS 1 (bases 1 to 99)
TITLE Haribabu,B. and Doltin,R.P.
JOURNAL Identification of a protein kinase multigene family of
MEDLINE Dictyostelium discoidium: Molecular cloning and expression of a
FEATURES cDNA encoding a developmentally regulated protein kinase
Proc. Natl. Acad. Sci. U.S.A. 88, 1115-1119 (1991)
91142122
source
1..99
/organism="Dictyostelium discoidium"
/strain="AX-3"
/db_xref="taxon:44689"
CDS
1..99
/codon_start=1
/product="protein kinase 5"
/protein_id="AA33390.1"
/db_xref="GI:167726"
/translation="NILLTSDGHIVLTDFGISKGLVSDNDRTATFC"
BASE COUNT 30 a 18 c 24 g 27 t
ORIGIN
Query Match 50.7%; Score 14.2; DB 34; Length 99;

Best Local Similarity 70.4%; Pred. No. 1.8e+04;
Matches 19; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 1 ccctatgttggtgccaaatggtca 27
||||| ||||| |||||
Db 60 CTCCTTGCTATACCAAAATCGTC 34

RESULT 15
E14994
LOCUS E14994 27 bp DNA PAT 28-JUL-1999
DEFINITION PCR primer to mutagenize equine cDNA encoding chorionic
gonadotropin alpha subunit at the 58th Asn to Gln.
ACCESSION E14994
VERSION E14994.1 GI:5709677
KEYWORDS JP 1998036285-A/8.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
AUTHORS 1 (bases 1 to 27)
TITLE Ogawa,T., Shioda,K., Bin,K. and Ikemi,M.
JOURNAL NEW GONADOTROPIC HORMONE AND ITS PRODUCTION
Patent: JP 1998036285-A 10-FEB-1998;
DENKI KAGAKU KOGYO KK
OS None
OC Artificial sequences.
PN JP 1998036285-A/8
PD 10-FEB-1998
PF 23-JUL-1996 JP 1996193232
PI OGAWA TOMOYA, SHIODA KUNIO, BIN KANSHYOKU, IKEMI MASAHIKA PC
A61K38/24,C07K14/59,C12N5/10,C12N15/09,C12P21/02,(C12P21/02, PC
C12R1:91);
CC strandedness: Single;
CC topology: Linear;
FH key
FH Location/Qualifiers
FT source 1..27
Location/Qualifiers
source
1..27
/organism="Artificial sequences"
/db_xref="taxon:32644"
BASE COUNT 8 a 8 c 5 g 6 t
ORIGIN

Query Match 50.0%; Score 14; DB 5; Length 27;
Best Local Similarity 77.3%; Pred. No. 2.1e+04;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 5 atgttggtgccaaatggtca 26
||||| ||||| |||||
Db 1 ATGTTGTCGCCAAAGACAGATCA 22
Search completed: June 4, 2000, 16:05:32
Job time: 27884 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:04 ; Search time 322.35 Seconds
(without alignments)
21.732 Million cell updates/sec

Title: US-09-164-714-8

Perfect score: 28
Sequence: 1 cccatgtgtgtgcacaaatgtgtcaag 28

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 31185 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	15.8	56.4	24	1	PCR primer for C.
C 2	15.2	54.3	29	1	5' PCR primer Deg-
C 3	15	53.6	34	1	Primer F1B2. Fibro
C 4	15	53.6	34	1	Primer for amplify
C 5	14.4	51.4	54	1	Oligo 1131 for hCG
C 6	14.4	51.4	60	1	Oligo 850 for hCG
C 7	14.4	51.4	84	1	Oligo 839 for hCG
C 8	14.2	50.7	72	1	Mutated complement
C 9	14.2	50.7	72	1	Mutant CDR oligonu
C 10	14.2	50.7	72	1	Mutant light chain
C 11	14	50.0	27	1	Equine chorionic g
C 12	13.8	49.3	35	1	Sequence of PCR pr
C 13	13.8	49.3	35	1	Sequence of PCR pr
C 14	13.8	49.3	67	1	Human gene signatu
C 15	13.8	49.3	71	1	Staphylococcus aur
C 16	13.6	48.6	21	1	PCR primer for ant
C 17	13.6	48.6	41	1	Human granulocyte
C 18	13.6	48.6	54	1	Staphylococcus aur
C 19	13.6	48.6	69	1	Staphylococcus aur
C 20	13.4	47.9	97	1	phc1sXVI mutant po
C 21	13.4	47.9	97	1	phc1sXVI mutant po
C 22	13.2	47.1	23	1	Human trkC recepto
C 23	13.2	47.1	26	1	Nucleotide sequenc
C 24	13.2	47.1	29	1	PCR primer CP19 to
C 25	13.2	47.1	29	1	PCR primer CP19 to
C 26	13.2	47.1	29	1	Canarypox ampliflc
C 27	13.2	47.1	29	1	Canarypox ampliflc
C 28	13.2	47.1	29	1	Canarypox primer.
C 29	13.2	47.1	29	1	Canarypox ppx4 PCR
C 30	13.2	47.1	29	1	Canarypox virus C3
C 31	13.2	47.1	29	1	Primer CP19 for 27
C 32	13.2	47.1	31	1	Human genomic DNA
C 33	13.2	47.1	35	1	Murine IgG PCRn ge
C 34	13.2	47.1	56	1	Synthetic fragment

C 35	13.2	47.1	59	1	Staphylococcus aur
C 36	13.2	47.1	72	1	Oligo used in synt
C 37	13.2	47.1	80	1	Ovine GHG primer O
C 38	13.2	47.1	94	1	Human gene signatu
C 39	13	46.4	21	1	Fragment encoding
C 40	13	46.4	24	1	Mumps haemagglutin
C 41	13	46.4	30	1	Phe-355 mutation o
C 42	13	46.4	34	1	Rev/tax mRNA of bo
C 43	13	46.4	48	1	Oligonucleotide pr
C 44	13	46.4	48	1	Test analyte from
C 45	13	46.4	55	1	PCR primer Al-7 us

ALIGNMENTS

RESULT 1	
X22061/c	
ID X22061 standard; DNA; 24 BP.	
AC X22061:	
DT 19-MAY-1999 (first entry)	
DE PCR primer for C. maltosa cytochrome POX4 gene.	
KW PCR primer: POX4 gene; Pichia pastoris; di-carboxylic acid production;	
KW 6-22C mono-carboxylic acid production; alkane hydroxylation; ss.	
OS Synthetic.	
OS Candida maltosa.	
PN W09904014-A2.	
PD 28-JAN-1999.	
PR 20-JUL-1998; U14935.	
PR 21-JUL-1997; US-053215.	
PA (DUPLO) DU POINT DE MEMOIRS & CO E. I.	
PI Fallon RD. Payne MS, Plicatagyo SK, Wu S;	
DR WPI: 99-132258/11.	
PT Production of mono- and di-carboxylic alkanic acids in new	
PT engineered yeast - containing inserted copies of genes for	
PT cytochrome P450 monooxygenase and/or reductase, providing increased	
PT production and conversion efficiency, useful e.g. as intermediates	
PT for polymers and surfactants	
PS Example 6; Page 33; 57pp; English.	
CC This sequence represents a PCR primer for the POX4 gene	
CC of Candida maltosa. The invention relates to a method for the	
CC production of 6-22C mono- and di-carboxylic acids (I), which comprises	
CC treating, under aerobic conditions: (i) Pichia pastoris having	
CC genetically engineered alkane hydroxylation activity; or (ii) Candida	
CC maltosa genetically engineered to have enhanced alkane hydroxylation	
CC activity and/or a blocked beta-oxidation pathway, with a 6-22C linear	
CC hydrocarbon. The method is specifically used to make dodecanedioic acid	
CC (Ia) from dodecane, but generally any 6-22C alkane, or derived	
CC monocarboxylic acids or esters, can be converted. (I) that are	
CC monocarboxylic acids are intermediates for surfactants and those that are	
CC dicarboxylic acids are intermediates for polymers and anticorrosion	
CC agents. Transformation with the specified genes introduces (I)-producing	
CC ability to P. pastoris and enhances that in C. maltosa, resulting in	
CC conversion efficiency and productivity suitable for a commercial process	
CC (that is less environmentally damaging than known chemical syntheses).	
CC Disruption of the beta-oxidation pathway increases metabolite flow to the	
CC omega-oxidation pathway, and thus yield of, and selectively for, (I).	
SO Sequence 24 BP: 6 A; 4 C; 4 G; 10 T;	
Query Match	56.4%; Score 15.8; DB 1; Length 24;
Best local Similarity	89.5%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Oy 9 tgggtccaaatgtgtca 27	
Db 23 tgggtccaaatgtgtca 5	
RESULT 2	
V43049/c	
ID V43049 standard; DNA; 29 BP.	
AC V43049:	

DT 23-OCT-1998 (first entry)
 DE 5' PCR primer Deg-A used to amplify nitric oxide synthase DNA.
 KW Nitric oxide synthase; treatment; epidermal; dermal condition;
 KM unwanted pigmentation; PCR primer; ss.
 OS Synthetic.
 OS Mus sp.
 OS Rattus sp.
 OS Homo sapiens.
 PN WO9833379-A1.
 PD 06-AUG-1998.
 PR 03-FEB-1998: U01891.
 PF 04-FEB-1997: US-037098.
 PA (GENO) GEN HOSPITAL CORP.
 PI Lerner EA, Qureshi AA;
 DR WPI: 98-437056/37.
 PT Treatment of unwanted epidermal or dermal conditions - comprising
 PT administration of treatment which modulates level of nitric oxide in
 PT skin
 PS Disclosure: Page 17; 35pp; English.
 CC PCR primers V43049-50 were used to amplify DNA encoding nitric oxide
 CC synthase (Inducible isoform, the isoform found in nerve cells, and
 CC isoform found in epithelial cells). The specification describes
 CC a method for treating a subject for an unwanted epidermal or dermal
 CC condition, which comprises administration of a compound which modulates
 CC the level of nitric oxide in the skin. Where the condition is
 CC characterised by unwanted cells or unwanted pigmentation, the method
 CC includes increasing the level of nitric oxide in the skin. Where the
 CC condition is characterised by the lack of or by an insufficient number
 CC of dermal or epidermal cells or a lack of pigmentation, the method
 CC includes reducing the level of nitric oxide in the skin.
 SQ Sequence 29 bp; 12 A; 7 C; 5 G; 2 T;

Query Match 54.3%; Score 15.2; DB 1; Length 29;
 Best Local Similarity 77.3%; Pred. No. 2.4e+02;
 Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ccctatgttggtgccaaattg 22
 |||| ||||||| | : ||||
 Db 27 CCTTGTGTGGTGCGCRACCTTG 6

RESULT 3
 ID Q71835 standard; DNA; 34 BP.
 AC Q71835;
 DT 30-MAR-1995 (first entry)
 DE Primer FIB2.
 KW Fibronectin binding protein; Fbp; fibronectin binding domain;
 KW monoclonal antibody; MAb; adhesion; Gram-positive bacterium;
 KW bacteremia; Escherichia coli; polymerase chain reaction; PCR;
 KW amplification; primer; Staphylococcus aureus; ss.
 OS Synthetic.
 PN WO9418327-A.
 PD 18-AUG-1994.
 PF 04-FEB-1994; G00215.
 PR 05-FEB-1993; GB-002289.
 PR 20-OCT-1993; GB-021592.
 PA (SMRK) SMITHKLINE BEECHAM PLC.
 PI Burnham MKR, Chopra I, Critchley IA, Knowles DJC;
 DR WPI: 94-279748/34.
 PT Fibronectin binding protein and monoclonal antibodies specific
 PT for Fbp - useful to prevent adherence of Gram-positive bacteria
 PT to indwelling devices or wounds
 PS Disclosure: Page 29; 40pp; English.
 CC Polypeptides corresponding to residues G709-T886 plus PPIVPT,
 CC G709-P838(P838) and G709-P838 (R58805-07, respectively) of
 CC S. aureus J2385 (NCIMB 40532) Fbp type A fibronectin binding
 CC domain D1-D4 region (R58808) were expressed in E. coli BL21(DE4)
 CC and used to raise MAbS specific for Fbp. DNA encoding
 CC D1-D4(G709-T886) was obtained by PCR amplification of S. aureus
 CC chromosomal DNA using primers FIB1 (Q71834) and FIB2 (Q71835).
 SQ Sequence 34 BP; 5 A; 6 C; 14 G; 9 T;

Query Match 53.6%; Score 15; DB 1; Length 34;
 Best Local Similarity 78.3%; Pred. No. 3e+02;
 Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ccctatgttggtgccaaattg 23
 || || ||||||| || |||||
 Db 7 CCTTACCTGTGGTGACAGATTGG 29

RESULT 4
 ID T18304 standard; DNA; 34 BP.
 AC T18304;
 DT 13-NOV-1996 (first entry)
 DE Primer for amplifying fibronectin binding domains.
 KW Fibronectin; oral hygiene; gingivitis; gum disease; adhesion;
 KW periodontitis; oral pathogen; caries; calculus; candidiasis;
 KW oral surgery; tissue regeneration; irrigation; toothpaste;
 KW dentifrice; mouthwash; lozenge; ss.
 OS Synthetic.
 PN WO9604003-A1.
 PD 15-FEB-1996.
 PF 18-JUL-1995; E02825.
 PR 05-AUG-1994; GB-015902.
 PA (SMRK) SMITHKLINE BEECHAM PLC.
 PI Barnett P, Critchley IA, Dodd I;
 DR WPI: 96-129122/13.
 PT Prevention of adherence of oral pathogens in the oral cavity,
 PT partic. tooth surfaces - by application of a fibronectin binding
 PT protein or polypeptide or a monoclonal antibody or fragment against
 PT it
 PS Example 1: Page 18; 41pp; English.
 CC Adherence of oral pathogens, particularly to tooth surfaces, can be
 CC prevented by application of a fibronectin binding protein or
 CC polypeptide. The fibronectin binding protein or polypeptide is
 CC useful in the manufacture of oral hygiene compositions, eg.
 CC toothpaste, liquid dentifrice, mouthwash or lozenge. They are
 CC useful to prevent the plaque-related development of carious lesions,
 CC gingivitis, calculus or periodontal disease and to combat oral cavity
 CC infections, e.g. candidiasis. They may also be useful in oral
 CC surgery e.g. in guided tissue regeneration procedures to prevent
 CC subsequent bacterial infection, and for irrigation of periodontal
 CC pockets. Two primers (T18303, T18304) were used to amplify DNA
 CC encoding the fibronectin binding domains of fibronectin binding
 CC protein from the chromosomal DNA of Staphylococcus aureus J2385.
 SQ Sequence 34 BP; 5 A; 6 C; 14 G; 9 T;

Query Match 53.6%; Score 15; DB 1; Length 34;
 Best Local Similarity 78.3%; Pred. No. 3e+02;
 Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ccctatgttggtgccaaattg 23
 || || ||||||| || |||||
 Db 7 CCTTACCTGTGGTGACAGATTGG 29

RESULT 5
 ID X27477 standard; DNA; 54 BP.
 AC X27477;
 DT 08-JUN-1999 (first entry)
 DE Oligo 1131 for hCG alpha-subunit analogues.
 KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
 KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
 KW human follicle stimulating hormone; human thyroid stimulating hormone;
 KW stability; primer; amplification; PCR; mutation; ss.
 OS Synthetic.
 OS Homo sapiens.
 PN WO9858957-A2.
 PD 30-DEC-1998.

PF 25-JUN-1998; U13070.
PR (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR:
DR WPI: 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Example 1; Page 60; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
CC mutants, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. Primers
CC X27450-X27484 were used to introduce the mutations in the hCG sequence.
CC The improved analogues are designed specifically to reduce perturbation
CC of the 3-dimensional structure of the hormone, thereby creating greater
CC likelihood that the dimer will be formed in vivo and the formed dimer
CC will have affinity for the native receptors and have agonistic activity
CC on them. The changes stabilise the GPHs and prolong the biological
CC activities of the hormones. The analogues can have uses as for the
CC native GPHs
SQ Sequence 54 BP; 18 A; 11 C; 12 G; 13 T;

Query Match 51.4%; Score 14.4; DB 1; Length 54;
Best Local Similarity 75.0%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 ctatgttggtgccaattgtca 26
|||||
DB 20 CTATGTGTGTCCTCAAAAGACGTC A 43

RESULT 6
X27467
ID X27467 standard; DNA; 60 BP.
AC X27467;
DT 08-JUN-1999 (first entry)
DE Oligo 850 for hCG alpha-subunit analogues.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation; ss.
OS Synthetic.
OS Homo sapiens.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR 25-JUN-1997; US-050784.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PI (MCIN-) MCINNIS P G.
PI Moyle WR:
DR WPI: 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Example 1; Page 60; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
CC mutants, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. Primers
CC X27450-X27484 were used to introduce the mutations in the hCG sequence.
CC The improved analogues are designed specifically to reduce perturbation
CC of the 3-dimensional structure of the hormone, thereby creating greater
CC likelihood that the dimer will be formed in vivo and the formed dimer
CC will have affinity for the native receptors and have agonistic activity
CC on them. The changes stabilise the GPHs and prolong the biological
CC activities of the hormones. The analogues can have uses as for the
CC native GPHs
SQ Sequence 84 BP; 25 A; 20 C; 16 G; 23 T;

Query Match 51.4%; Score 14.4; DB 1; Length 84;
Best Local Similarity 75.0%; Pred. No. 6.3e+02;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 ctatgttggtgccaattgtca 26
|||||
DB 44 CTATGTGTGTCCTCAAAAGACGTC A 67

RESULT 8
O97403/c
ID O97403 standard; DNA; 72 BP.
AC O97403;

CC likelihood that the dimer will be formed in vivo and the formed dimer
CC will have affinity for the native receptors and have agonistic activity
CC on them. The changes stabilise the GPHs and prolong the biological
CC activities of the hormones. The analogues can have uses as for the
CC native GPHs.
SQ Sequence 60 BP; 18 A; 13 C; 14 G; 15 T;

Query Match 51.4%; Score 14.4; DB 1; Length 60;
Best Local Similarity 75.0%; Pred. No. 6e+02;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 ctatgttggtgccaattgtca 26
|||||
DB 20 CTATGTGTGTCCTCAAAAGACGTC A 43

RESULT 7
X27464
ID X27464 standard; DNA; 84 BP.
AC X27464;
DT 08-JUN-1999 (first entry)
DE Oligo 839 for hCG alpha-subunit analogues.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation; ss.
OS Synthetic.
OS Homo sapiens.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR 25-JUN-1997; US-050784.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PI (MCIN-) MCINNIS P G.
PI Moyle WR:
DR WPI: 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Example 1; Page 60; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
CC mutants, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. Primers
CC X27450-X27484 were used to introduce the mutations in the hCG sequence.
CC The improved analogues are designed specifically to reduce perturbation
CC of the 3-dimensional structure of the hormone, thereby creating greater
CC likelihood that the dimer will be formed in vivo and the formed dimer
CC will have affinity for the native receptors and have agonistic activity
CC on them. The changes stabilise the GPHs and prolong the biological
CC activities of the hormones. The analogues can have uses as for the
CC native GPHs.
SQ Sequence 84 BP; 25 A; 20 C; 16 G; 23 T;

Query Match 51.4%; Score 14.4; DB 1; Length 84;
Best Local Similarity 75.0%; Pred. No. 6.3e+02;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 ctatgttggtgccaattgtca 26
|||||
DB 44 CTATGTGTGTCCTCAAAAGACGTC A 67

RESULT 8
O97403/c
ID O97403 standard; DNA; 72 BP.
AC O97403;

DE	01-APR-1996 (first entry)
DT	Mutated complementarity determining region L1.
KM	Polymerase chain reaction; PCR; primer; amplifi;
KW	chimeric sequence; interleukin-1; antibody; tpa;
RN	repetitive DNA; ribozyme; aptamer; gene therapy;
KM	complementarity determining region; ss.
OS	Synthetic.
FH	Key
FH	Location/Qualifiers
FT	misc_difference 22. .51
FT	/tag= A
FT	/note= "mutagenised region of wild type CDR H2 (see
FT	Q96186 for wild type)."
PN	WO9523625-A1.
PD	24-AUG-1995.
PF	17-FEB-1995; U02126.
PR	17-FEB-1994; US-198431.
PA	(AFFY-) AFFYMAX TECHNOLOGIES NV.
PI	Cramerl A. Stemmer WPC;
DR	WPt: 95-302727/39.
PT	DNA mutagenesis via random fragmentation and reassembly - useful in
PT	the produ. of mutant proteins having enhanced biological activity
PS	Example 7; Page 82; 120pp; English.
CC	The sequences shown in Q97400-Q97405 are mutated versions of the
CC	complementarily determining regions (CDRs) shown in Q96183-Q96188. These
CC	mutated CDRs were added to reassembly products of the A10B wild-type
CC	antibody gene. The reassembled products were created by allowing the
CC	50-200 bp fragments, produced upon DNaseI digestion, to reassemble. Once
CC	the CDRs were added shuffling was allowed to occur. This method can also
CC	be used to shuffle whole gene sequences. The method can also be used to
CC	produce recombinant proteins, and chimeric DNA sequences. The methods
CC	can be used to produce proteins with a desired phenotype having an
CC	advantageous predetermined selectable property. Proteins such as
CC	interleukin-1, antibodies, tpa and growth hormones can be generated that
CC	have altered specificity or activity. Mutant sequences such as promoter
CC	regions, introns, exons and enhancer sequences can also be generated,
CC	which can be used to generate genes having increased expression rates.
CC	The method can be used in the study of repetitive DNA regions, and for
CC	the mutation of ribozymes or aptamers. Shifted viral proteins which may
CC	comprise epitopes that will arise by viral evolution can also be
CC	produced. This method of shuffling sequences may be used for generating
CC	gene therapy vectors and replication-defective gene therapy constructs.
CC	The advantage with this method is that it is not necessary to know the
CC	actual DNA or RNA sequence of the specific nucleic acid fragment.
SO	Sequence 72 BP; 21 A; 17 C; 18 G; 16 T;
Query Match	50.7%; Score 14.2; DB 1; Length 72;
Best Local Similarity	70.4%; Pred. No. 7.6e+02;
Matches 19; Conservative 0; Mismatches 8; Indels 0; Gaps 0.	
Oy	2 cctatgttgatgccaaattggtcacaag 28
Db	65 CTTCTGTGGTACCAATATATCTAACG 39
RESULT 9	
T73061/c	T73061 standard; DNA: 72 BP.
AC	T73061:
DT	13-FEB-1998 (first entry)
DE	Mutant CDR oligonucleotide L1 for the murine antibody A10B scFv gene.
KM	TEM-1 betalactamase gene; gene reassembly reaction; DNA library;
KM	gene shuffling; random fragmentation; mutagenesis; recombination;
KW	GFP protein; arsenate detoxification bacteria; cetotaxime; antibody A10B;
KM	cadmium detoxification bacteria; drug resistance gene; ss.
OS	Synthetic.
PN	WO9720078-A1.
PD	05-JUN-1997.
PF	02-DEC-1996; U19256.
PR	25-MAR-1996; US-621859.
RA	30-NOV-1995; US-564955.
PA	(AFFY-) AFFYMAX TECHNOLOGIES NV.
PI	Cramerl A. Stemmer WPC;

DR WP1: 97-310638/28.
 PR Generating polynucleotide(s) with desired characteristics by
 PR iterative selection and recombination - used for the directed
 PR molecular evolution in vitro or in vivo of proteins, especially
 PR green fluorescent protein
 PS Example 7: Page 127, 209pp; English.
 CC Sequences T73058-62 represent mutant complementarity determining regions
 CC (CDRs) of the A10B scv murine antibody gene. The present sequence
 CC represents a CDR in the light chain. The original A10B antibody
 CC reproducibility had only a low avidity, as it bound weakly to antigen. A
 CC novel method was used to improve the antibody by DNA shuffling of a
 CC library of all 6 mutant CDRs (based on CDRs T73046-51). The method
 CC involves recombining at least 2 forms of a polynucleotide to
 CC produce a library of recombinant forms of the sequence. The library is
 CC screened for a recombinant sequence, which is then recombined with
 CC another form of the polynucleotide, the same or different from the first
 CC and second forms, to produce a further library of recombinant
 CC polynucleotides. The subsequent libraries are screened, and recombination
 CC performed until the recombinant polynucleotide has acquired the desired
 CC property. The methods are used to reassemble DNA after random
 CC fragmentation, for mutagenesis of nucleic acid sequences by in vitro or
 CC in vivo recombination. The repeated cycles of mutagenesis, shuffling and
 CC selection allow for the directed molecular evolution in vitro or in vivo
 CC of proteins. In particular the methods are used for the production of
 CC mutant GFP protein which has been modified to show enhanced fluorescence
 CC when used as a reporter of gene expression and regulation. Also
 CC exemplified is the use of the methods to generate improved arsenate
 CC or cadmium detoxification bacteria, drug resistance genes and variant
 CC polymerases.
 SQ Sequence 72 BP; 21 A; 17 C; 18 G; 16 T;

Query Match 50.7%; Score 14.2; DB 1; Length 72;
 Best Local Similarity 70.4%; Pred. No. 7.6e+02;
 Matches 19; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 cctatgttgatgcgaattggtcaag 28
 | | | | | | | | | | | | | | | | | | | |
 DB 65 cttctgttgatgcgaattatgtatcag 39

RESULT 10
 VS4920/c
 ID VS4920 standard; DNA; 72 BP.
 AC VS4920;
 DT 19-NOV-1998 (first entry)
 DE Mutant light chain CDR L1 sequence.
 DE A10B antibody gene; recombination; gene shuffling; ss.
 OS Synthetic.
 PN US5811238-A.
 PD 22-SEP-1998.
 PF 30-NOV-1995; 564955.
 PR 04-MAR-1996; US-537874.
 PR 17-FEB-1994; US-198431.
 PR 30-NOV-1995; US-564955.
 PA (AFRY-) AFRYMAX TECHNOLOGIES NV.
 PI Cramerit A, Stemmer WPC;
 DR WP1: 98-530860/45.
 PR Production of polynucleotides with desired properties - by iterative
 PR selection and recombination
 PS Example 7: Columns 55-56; 74pp; English.
 CC The present sequence represents a mutant complementarity determining
 CC region (CDR) L1 (see VS4926 for original sequence) from the light chain
 CC of the A10B antibody gene. The antibody was improved by DNA shuffling of
 CC a library of 6 mutant CDRs, using the method of the invention. The
 CC specification describes a method for evolving a polynucleotide for
 CC application of a desired property. The method comprises providing a
 CC population of variants of the polynucleotide, at least one of which is
 CC in cell-free form, shuffling the variants of the polynucleotide to form
 CC recombinant polynucleotides, selecting or screening for recombinant
 CC polynucleotides that have evolved toward the desired property and
 CC repeating the steps with the selected recombinant polynucleotides until
 CC a recombinant polynucleotide has acquired the desired property. The

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:05 ; Search time 4521.53 Seconds
(without alignments)
25.100 Million cell updates/sec

Title: US-09-164-714-8

Perfect score: 28

Sequence: 1 ccctatgtgtgcacaaattgtcaag 28

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST.*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result

Query

No.	Score	Match	Length	DB	ID	Description
1	15.6	55.7	67	34	AA474067	AA474067 ves4e02.x
2	15.2	54.3	46	43	AI209047	AI209047 qg28e02.x
3	15	53.6	73	51	AI739189	AI739189 w127c11.x
4	15	53.6	82	51	AI745636	AI745636 tr04g11.x
5	15	53.6	89	23	H44598	H44598 ypl7c09.t1
6	15	53.6	93	44	AI282700	AI282700 qf83h07.x
7	15	53.6	97	37	AA714476	AA714476 nw06f12.s
8	14.8	52.9	97	35	AA566933	AA566933 987 l0bio
9	14.8	52.9	87	60	AI795116	AI795116 sb7b05.y
10	14.6	52.1	55	46	AI444423	AI444423 fb2b06.x
11	14.6	52.1	59	62	AI880479	AI880479 at78d09.x
12	14.6	52.1	92	23	H40392	H40392 ype0e01.t1
13	14.6	52.1	99	70	AA145122	AA145122 ga29e08.y
14	14.4	51.4	79	29	AA165762	AA165762 ms60f11.x
15	14.4	51.4	91	81	AA394554	AA394554 sb32e03.y
16	14.4	51.4	96	44	AI314653	AI314653 u127a11.x
17	14.4	51.4	98	41	AU008426	AU008426 AU008426
18	14.2	50.7	98	28	AA107577	AA107577 mp05a11.r
19	14.2	50.7	97	29	AA180917	AA180917 zp44b01.r
20	14	50.0	46	40	AA995439	AA995439 os84f01.s
21	14	50.0	88	47	AI474685	AI474685 tm35f12.x
22	14	50.0	91	36	AA659101	AA659101 nu81e01.s
23	13.8	49.3	86	29	AA146242	AA146242 mq88f06.r
24	13.8	49.3	89	26	W76217	W76217 zd58f09.t1
25	13.8	49.3	92	46	AI423670	AI423670 tf85c04.x
26	13.8	49.3	99	41	AU007058	AU007058 AU007058
27	13.6	48.6	56	84	B03327	B03327 CSRL-176D2-
28	13.6	48.6	61	38	AA746401	AA746401 nw62c10.s
29	13.6	48.6	61	38	AI338554	AI338554 qg93c08.x
30	13.6	48.6	88	24	N25181	N25181 yy02b10.s1
31	13.6	48.6	91	34	AA472378	AA472378 vho6g06.r
32	13.6	48.6	95	40	AA910572	AA910572 ok66h10.s
33	13.6	48.6	97	44	AI285065	AI285065 qk56f09.x
34	13.6	48.6	98	29	AA187895	AA187895 zp74g12.r
35	13.6	48.6	98	50	F29909	F29909 HSPD20053.H
36	13.4	47.9	58	41	AU007255	AU007255 AU007255
37	13.4	47.9	64	46	AI442743	AI442743 sa85h01.y
38	13.4	47.9	67	30	AA232723	AA232723 z775e10.r
39	13.4	47.9	67	37	AA669910	AA669910 ag42g07.s
40	13.4	47.9	68	37	AA667085	AA667085 vr87c12.s
41	13.4	47.9	79	33	AA446093	AA446093 zw58a10.s
42	13.4	47.9	98	46	AI427907	AI427907 m138e10.x
43	13.4	47.9	100	29	AA152082	AA152082 z148g12.r
44	13.2	47.1	34	45	AI367088	AI367088 qg46d03.x
45	13.2	47.1	49	24	H97554	H97554 yw04h01.s1

ALIGNMENTS

RESULT 1
 LOCUS AA474067 67 bp mRNA EST 18-JUN-1997
 DEFINITION ves4e02.r1 Beddington mouse embryonic region Mus musculus cDNA
 clone IMAGE:821978 5' similar to TR:G476095 G476095 B4-2 PROTEIN.
 ; RNA sequence.

ACCESSION AA474067
 VERSION AA474067.1 GI:22022294
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 67)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisell,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Willson,R. and
 Waterston R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)

COMMENT

On May 9, 1995 this sequence version replaced gi:803068.
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LBNL; contact the
 IMAGE Consortium (info@lmbl.lbl.gov) for further information.
 MGI:490258
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28ml3 rev1 ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES

source
 1..67
 /organism="Mus musculus"
 /strain="C57BL6 x DBA"
 /db_xref="taxon:10090"
 /clone="IMAGE:821978"
 /clone_lib="Beddington mouse embryonic region"
 /sex="pooled"
 /tissue_type="embryo"
 /dev_stage="7.5dpc"
 /lab_host="DH12S"
 /note="Organ: whole embryo; Vector: pCMV-SPORT; Site:1:
 SalI; Site:2: NotI; Cloned unidirectionally. Primer:
 oligo dt. Gastrulating embryos were collected at 7.5dpc
 from C57BL6 x DBA matings, excluding embryos that had
 developed head folds and all extraembryonic tissues.
 Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
 Referenced in Development 121, 2479-2489 (1995)"
 BASE COUNT
 19 a 23 c 13 g 12 t

Query Match

Best Local Similarity 81.8%; Score 15.6; DB 34; Length 67;
 Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 ctatgtgtgtgcgaattggt 24
 ||||| ||||| ||||| |||||
 Db 4 CTATGCTGGAGCCAAATTAGT 25

RESULT 2
 LOCUS AI209047 46 bp mRNA EST 29-NOV-1998
 DEFINITION qg28e02.x1 NCI_CGAP_K1d3 Homo sapiens cDNA clone IMAGE:1762394 3'

similar to SM:Y124_HUMAN Q14137 HYPOPHYRETICAL PROTEIN KIA0124 ;
 RNA sequence.

ACCESSION AI209047
 VERSION AI209047.1 GI:3770989
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 46)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT On Jan 19, 1998 this sequence version replaced gi:2151684.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Koskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbdp/image/image.html

Trace considered overall poor quality
Insert Length: 692 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES

source

1. .46

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1762394"

/lab_host="NCI-CGAP_K1d3"

/lab_host="DH10B"

/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I - oligo(dT) primer, strand cDNA was primed with a Not I - oligo(dT) primer, double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. mRNA source: 2 pooled kidneys. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 10 a 9 c 17 g 10 t

ORIGIN

Query Match

Best Local Similarity 54.3%; Score 15.2; DB 43; Length 46;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 tgggtcccaaatcgtcgaag 28

||||| ||||| ||||| |||||

Db 18 TCGTGCMAAAGTGTGAAG 37

RESULT 3

AI739189

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 73)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On Jun 5, 1998 this sequence version replaced gi:3189015.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Ian Kirsch, M.D., Michael R. Emmert-Buck,

M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/dbdp/image/image.html

FEATURES

source

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. .73

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2391476"

/clone_1lb="NCI-CGAP_CO16"

/tissue_type="colon tumor, RER"

/lab_host="DH10B"

/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP_CO10 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneds 1057416-1061255, and 114584-1145351).

Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 17 a 18 c 13 g 25 t

ORIGIN

Query Match

Best Local Similarity 53.6%; Score 15; DB 51; Length 73;

Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 cccatgttggtcccaaatg 23

||||| ||||| ||||| |||||

Db 36 CCCATGTGTGTGACACGCTGG 58

RESULT 4

AI745436

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 82)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On May 18, 1998 this sequence version replaced gi:3137118.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/dbdp/image/image.html

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. .82

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2217380"

/clone_1lb="NCI-CGAP_OV23"

/tissue_type="tumor, 5 pooled (see description)"

/lab_host="DH10B"

/note="Organ: ovary; Vector: PCMV-SPORE; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.35 kb. Tumor types include: mixed

Mullerian tumor, papillary serous, clear cell, spindle

cell. All are primary tumors, metastasis positive. Life

Technologies catalog #: 11534-013"

BASE COUNT 17 a 15 c 10 g 40 t

ORIGIN

Query Match 53.6%; Score 15; DB 51; Length 82;
Best Local Similarity 78.3%; Pred. No. 6e+03;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 cccatgtgtgtgtgccaattgg 23
|||||
Db 54 CCTATGTTGTAGCCACAGCTGG 76

RESULT 5

LOCUS H44598 89 bp mRNA EST 31-JUL-1995
DEFINITION ypl7c09.r1 Soares breast 3BDHbst Homo sapiens cDNA clone
IMAGE:187696 5' similar to gb:W80563 PLACENTAL CALCIUM-BINDING
PROTEIN (HUMAN);, mRNA sequence.

ACCESSION H44598
VERSION H44598.1 GI:920650
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 89)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

Insert Size: 585
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Trace considered overall poor quality
Insert length: 585 Std Error: 0.00
Seq primer: M13Rev
High quality sequence stop: 1.
Location/Qualifiers

FEATURES

source

1. 89
/organism="Homo sapiens"
/db_xref="GDB:3818592"
/db_xref="taxon:9606"
/clone="IMAGE:187696"
/clone_lib="Soares breast 3BDHbst"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: breast; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - Oligo(dT) primer [5'
TGTTCACATCTGAGTGGAGCGCGCCCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of a modified pT73 vector (Pharmacia).
Library went through one round of normalization to a Cot =
20. Library constructed by Bento Soares and M. Fatima
Bonaldo."

BASE COUNT 22 a 19 c 15 g 23 t 10 others
ORIGIN

Query Match 53.6%; Score 15; DB 23; Length 89;
Best Local Similarity 72.0%; Pred. No. 6.1e+03;
Matches 18; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 cccatgtgtgtgccaattgtca 26
|||||
Db 38 CCTTGTGTGNTGNCAGAGTGTCTAA 14

RESULT 6

LOCUS A1282700 93 bp mRNA EST 23-NOV-1998
DEFINITION qc83h07.x1 NCI-CGAP_Col4 Homo sapiens cDNA clone IMAGE:1961917 3'
similar to gb:X17206 40S RIBOSOMAL PROTEIN S4 (HUMAN);, mRNA
sequence.

ACCESSION A1282700
VERSION A1282700.1 GI:3920933
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 93)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

COMMENT On Jan 17, 1998 this sequence version replaced gi:1900073.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/dbp/image/image.html

Seq primer: -40UP from Gibco.
Location/Qualifiers

FEATURES

source

1. 93
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1961917"
/clone_lib="NCI-CGAP_Col4"
/tissue_type="moderately-differentiated adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: PCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.7 kb. Life Technologies catalog #:
11531-019"

BASE COUNT 20 a 22 c 21 g 30 t
ORIGIN

Query Match 53.6%; Score 15; DB 44; Length 93;
Best Local Similarity 78.3%; Pred. No. 6.1e+03;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 cccatgtgtgtgtgccaattgg 23
|||||
Db 37 CCTATGTTGTAGCCACAGCTGG 59

RESULT 7

LOCUS AA714476 97 bp mRNA EST 22-JAN-1998
DEFINITION nw06f12.s1 NCI-CGAP_S51 Homo sapiens cDNA clone IMAGE:1238543 3'
similar to gb:X17206 40S RIBOSOMAL PROTEIN S4 (HUMAN);, mRNA
sequence.
ACCESSION AA714476


```

VERSION      AA/14476.1  GI:2726750
KEYWORDS
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS      Eukarya; Primates; Catarrhini; Hominiidae; Homo.
TITLE        1 (bases 1 to 97)
REFERENCE    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE        Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      On Sep 12, 1996 this sequence version replaced gi:1407494.
              Contact: Robert Strausberg, Ph.D.
              Tel: (301) 496-1550
              Email: Robert.Strausberg@nih.gov
              Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
              Emmert-Buck, M.D., Ph.D.
              CDNA Library Preparation: Stratagene, Inc.
              CDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LINL at:
              www.bio.lnl.gov/db/brp/image/image.html

Insert Length: 771      Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1238543"
/clone_lib="NCI-CGAP-S61"
/tissue_type="synovial sarcoma"
/lab_host="SOLR (kanamycin resistant)"
/Note="Vector: Bluescript SK-; Site.1: EcoRI; Site.2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. Synovial
sarcoma. 5' adaptor sequence: 5' GAATTCGGCAG 3' 3'
adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' Average
insert size: 1.4 kb."

BASE COUNT   18 a      24 c      26 g      29 t
ORIGIN
Query Match          53.6%; Score 15; DB 37; Length 97;
Best Local Similarity 78.3%; Pred. No. 6.2e+03;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 cccatgttggtgccaaattgg 23
    |||||||  |||  |  |||
Db 25 cccatgttggtgacccacgctgg 47

RESULT 8
AA566933 59 bp mRNA EST 28-AUG-1998
LOCUS 987 lobiolly pine CA pinus taeda cDNA clone 1CA12G, mRNA sequence.
ACCESSION AA566933
VERSION AA566933.1 GI:3366146
KEYWORDS EST.
SOURCE lobiolly pine.
ORGANISM Pinus taeda
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
TITLE Pinaceae; Pinus.
JOURNAL 1 (bases 1 to 59)
AUTHORS Allona, I., Quinn, M., Shoop, E., Swope, K., St. Cyr, S., Carlis, J.,
TITLE Riedl, J., Retzel, E., Campbell, M., Sederoff, R. and Whetten, R.W.
JOURNAL Analysis of xylem formation in pine by cDNA sequencing
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 95 (16), 9693-9698 (1998)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394808.
Contact: Ross Whetten

```

```

FOREST Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall,
Raleigh, NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhet@unc1y.ncsu.edu
Seq primer: T3.

FEATURES
source
Location/Qualifiers
1. 59
/organism="Pinus taeda"
/straln="Coastal plain lobiolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="1CA12G"
/clone_lib="lobiolly pine CA"
/tissue_type="Xylem"
/lab_host="SOLR"
/Note="Vector: lambda-ZAP; Site.1: EcoRI; Site.2: XhoI;
The result of subtraction of C library with N library.
Immature xylem from the underside of inclined stems of
differentiating compression wood was subtracted with
immature xylem from the side of inclined stems of
differentiating wood. A mixture of four genotypes were
used. Oligo-dT primed cDNA was directionally cloned into
the EcoRI-XhoI lambda-ZAP vector arms"

BASE COUNT   15 a      12 c      16 g      14 t
ORIGIN
Query Match          52.9%; Score 14.8; DB 35; Length 59;
Best Local Similarity 84.2%; Pred. No. 7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 tgggtccaaattggtcaa 27
    |||||||  |||  |  |||
Db 14 TGGTGCCANATGATCA 32

RESULT 9
A1795116 87 bp mRNA EST 13-DEC-1999
LOCUS sb76h05.y1 Gm-cl010 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl010-946 5', mRNA sequence.
ACCESSION A1795116
VERSION A1795116.1 GI:5342832
KEYWORDS EST.
ORGANISM soybean.
REFERENCE Glycine max
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
JOURNAL eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
COMMENT Glycine.
1 (bases 1 to 87)
Shoemaker, R., Keim, P., Vodkin, L., Erpelidg, J., Coryell, V.,
Khanan, A., Bolla, B., Maira, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
On Jun 22, 1998 this sequence version replaced gi:3247195.
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or

```

info@genomesystems.com web site: www.genomesystems.com.

FEATURES
source
Location/Qualifiers
1. .87
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1010-946"
/tissue_type="young cotyledons of greenhouse grown plants"
/dev_stage="2cm long 12 week old"
/lab_host="X110-Gold"
/note="Vector: pBluescript II SK(+). Site 1: EcoRI; Site 2: XhoI. This cDNA library was constructed from mRNA isolated from immature cotyledons (100-200mg) of old greenhouse grown plants. The cDNA library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into X110-gold host cells. This library was constructed by Dr. Ilya Vodkin and Dr. Anu Khanna."

BASE COUNT
ORIGIN
17 a 15 c 19 g 36 t

Query Match
Best Local Similarity 52.9%; Score 14.8; DB 60; Length 87;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ctatgttggtgcccaaat 20
11 |||||||||||
Db 41 CTGTGTGTGTCCTCAACT 58

RESULT 10
LOCUS A1444123 55 bp mRNA EST 09-MAR-1999
DEFINITION fb26b06.x1 zebrafish Washu MPING EST Danio rerio cDNA 3' similar to gb:M80254 PEPTIDYL-PROLYL CIS-TRANS ISOMERASE, MITOCHONDRIAL
PRECUSOR (HUMAN);, mRNA sequence.
ACCESSION A1444123
VERSION A1444123.1 GI:4307970
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Chordata; Craniata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
1 (bases 1 to 55)
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowles,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
Washu zebrafish EST Project 1998
Unpublished (1998)
On Mar 20, 1998 this sequence version replaced gi:2980362.
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrfish@wustl.edu
CDNA Library Preparation: Matthew Clark. CDNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenzentrumPrimaDatenbank, Berlin, Germany (web address:

www.rzpd.de)
Trace considered overall poor quality
Seq primer: 17 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .55
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="1b-"zebrafish Washu MPING EST"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield stage embryos"
/lab_host="X11-blue MRP"
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; 1st strand cDNA was primed with a Not I - Oligo(dT)15 primer [5'pGACGTGCTTAGATGCGAGCGCGCCGCTTTTCTTTTCTT3']; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab; ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

BASE COUNT
ORIGIN
12 a 10 c 19 g 14 t

Query Match
Best Local Similarity 52.1%; Score 14.6; DB 46; Length 55;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 6 tgttggtgcccaattgtca 26
||||||| ||| ||
Db 27 TGTTGTGTCGACATTGGCA 47

RESULT 11
LOCUS A1880479 59 bp mRNA EST 23-AUG-1999
DEFINITION at78d09.x1 Barsstead colon HPLR87 Homo sapiens cDNA clone IMAGE:2378129 3' similar to gb:X17206 40S RIBOSOMAL PROTEIN S4 (HUMAN);, mRNA sequence.
ACCESSION A1880479
VERSION A1880479.1 GI:5554528
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 59)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
Washu-NCI human EST Project
Unpublished (1997)
On Mar 10, 1998 this sequence version replaced gi:2948486.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

/lab_host="DH10B"
/note="Vector: pBluescript SK-, Site_1: EcoRI; Site_2: XhoI; Construction of the cDNA library was carried out using Stratagene's 'unizap - cDNA synthesis kit'. cDNA was constructed using an oligo dt primer/linker that contains a XhoI site within it. Following ds cDNA synthesis, EcoRI adapters were ligated to the blunt ends and sample was digested with XhoI. The result is cDNA with an EcoRI sticky end on one side and a XhoI sticky end on the other. This cDNA was ligated directionally in unizap arms. The vector is designed containing the pBluescript sequence as well as lambda DNA and cDNA is cloned within this pBluescript sequence. The vector was then packaged using gold giga packaging extracts. Library was grown in XLBlue MRF cells and amplified. The library was excised by mass excision using Stratagene's 'Mass excision kit' that uses exasist as a helper phage that releases the pBluescript sequence and circularizes it as single stranded plasmids that are then packaged (by helper phage) and secreted out of the host cell as phagemids. SOLR cells were transformed with phagemids and the library was plated out on LB-amp plates to select for transformants. Approximately 1,000,000 colonies were grown and recovered. The double stranded plasmid library was recovered by using Qiaagen Midi prep kit. 2 micro grams of each library were used to transform DH10B cells by electroporation."

BASE COUNT
ORIGIN
27 a 23 c 24 g 25 t

Query Match 52.1%; Score 14.6; DB 70; Length 99;
Best Local Similarity 81.0%; Pred. No. 9.4e+03;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 8 ttgtgtccaaatgtgtcaag 28
||||| ||||| ||||| |||||
DB 89 TTGGACCGCAATTGGACCAAG 69

RESULT 14
AA165762/c
LOCUS
DEFINITION
AA165762 79 bp mRNA EST 12-FEB-1997
mus60f11.r1 Stratagene mouse embryonic carcinoma (8337317) Mus
musculus cDNA IMAGE:615961 5' similar to TR:E93245 E93245 ETN
INSERT IN THE FAS APOPTOSIS GENE OF MRL-LPR/IPR. [1] ; mRNA
sequence.
ACCESSION
KEYWORDS
AA165762
AA165762.1 GI:1743977
EST.
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 79)
Marras,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE
JOURNAL
COMMENT
The Washu-HMT Mouse EST Project
Unpublished (1996)
On Sep 12, 1996 this sequence version replaced gi:1325676.
Contact: Maira M/Mouse EST Project
Washu-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1376805
Seq primer: -28m13 rev1 ET from Amersham

FEATURES
source
High quality sequence stop: 75.
Location/Qualifiers
1..79
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:615961"
/clone_1ib="Stratagene mouse embryonic carcinoma
(#937317)"
/tissue_type="carcinoma"
/dev_stage="embryonic"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-, Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. P19 cell
line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGACAGAG 3' -3' adaptor
sequence: 5' CTCGACTTTTGTTTTGTTTT 3' "

BASE COUNT
ORIGIN
20 a 19 c 26 g 14 t

Query Match 51.4%; Score 14.4; DB 29; Length 79;
Best Local Similarity 75.0%; Pred. No. 1.1e+04;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
OY 1 ccctatgttgcgaatgtgt 24
||||| ||||| ||||| |||||
DB 56 CCCAGTTCGCGCCCAAAATGTGT 33

RESULT 15
AW394554
LOCUS
DEFINITION
AW394554 91 bp mRNA EST 07-FEB-2000
sh32e03.y1 Gm-c1017 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1017-3413 5', mRNA sequence.
ACCESSION
KEYWORDS
AW394554
AW394554.1 GI:6912966
EST.
SOURCE
ORGANISM
soybean.
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eudicots I; Fabales; Fabaceae; Papilionoideae;
Glycine.
1 (bases 1 to 91)
Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Corryell,V.,
Khanna,A., Bolla,B., Marras,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurr,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
On Jul 9, 1999 this sequence version replaced gi:5433560.
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Seq primer: -40RP from Gibco.
Location/Qualifiers
1..91
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1017-3413"
/clone_1ib="Gm-c1017"
/tissue_type="vegetable buds of field grown plants"

```

/lab_host="XL10-Gold"
/Note="Vector: pbluescript II XR: Site_1: EcoRI: Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from vegetable buds of field grown plants. The cDNA
library was prepared using the Stratagene pbluescript II
XR library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(AT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pbluescript vector. The ligated cDNA fragments were
transformed into XL10-Gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John
Eipelidng."

```

```

BASE COUNT      22 a      33 g      25 t
ORIGIN          11 c

```

```

Query Match      51.4%; Score 14.4; DB 81; Length 91:
Best Local Similarity 75.0%; Pred. No. 1.1e+04;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY      1 ccctatgtgtgtccaaatgtgt 24
        ||||| ||||| |||||
Db      44 CCCTTGTGTGTAATAATGCTT 67

```

```

Search completed: June 4, 2000, 13:53:09
Job time: 20933 sec

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:40 ; Search time 244.64 Seconds
(without alignments)
14.877 Million cell updates/sec

Title: US-09-164-714-8

Perfect score: 28

Sequence: 1 cccatgtgtgtgcacaaatgtgtcaag 28

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/2/lna/5C.COMB.seq:*
4: /cgn2_6/prodata/2/lna/5D.COMB.seq:*
5: /cgn2_6/prodata/2/lna/6.COMB.seq:*
6: /cgn2_6/prodata/2/lna/PCRUS.COMB.seq:*
7: /cgn2_6/prodata/2/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	57.1	39	1	US-08-361-920-32
2	16	57.1	39	1	US-08-479-939-32
3	16	57.1	39	2	US-08-483-432-32
4	15	53.6	34	4	US-08-459-135A-2
5	14.2	50.7	72	1	US-08-198-431-31
6	14.2	50.7	72	2	US-08-564-955-31
7	14.2	50.7	72	3	US-08-537-874-31
8	14	50.0	36	2	US-08-764-100-6
9	13.8	49.3	35	3	US-08-583-276-9
10	13.8	49.3	35	3	US-08-583-276-10
11	13.6	48.6	21	4	US-08-743-637B-210
12	13.6	48.6	26	2	US-08-538-875-26
13	13.4	47.9	18	5	US-09-289-376-21
14	13.4	47.9	44	2	US-08-592-406-5
15	13.4	47.9	97	1	US-07-635-561A-3
16	13.2	47.1	23	3	US-08-359-705B-11
17	13.2	47.1	23	3	US-08-286-846A-11
18	13.2	47.1	23	4	US-08-457-880A-11
19	13.2	47.1	23	5	US-08-444-622A-11
20	13.2	47.1	23	5	US-08-942-562-11
21	13.2	47.1	29	1	US-08-105-483-268
22	13.2	47.1	29	2	US-08-224-391-84
23	13.2	47.1	29	2	US-08-484-304-84
24	13.2	47.1	29	2	US-08-224-657-105
25	13.2	47.1	29	2	US-08-709-209-268
26	13.2	47.1	29	2	US-08-257-073-63
27	13.2	47.1	29	2	US-08-458-101-268

c 28	13.2	47.1	29	3	US-08-184-009-134	Sequence 134, App
c 29	13.2	47.1	29	3	US-08-566-398-45	Sequence 45, App
c 30	13.2	47.1	29	4	US-08-458-356-134	Sequence 134, App
c 31	13.2	47.1	29	4	US-08-658-665-88	Sequence 88, App
c 32	13.2	47.1	43	1	US-07-885-689A-24	Sequence 24, App
c 33	13.2	47.1	43	4	US-08-857-946-116	Sequence 116, App
c 34	13.2	47.1	43	5	US-08-970-740-116	Sequence 116, App
c 35	13	46.4	21	1	US-08-434-411-38	Sequence 38, App
c 36	13	46.4	21	2	US-08-434-402-38	Sequence 38, App
c 37	13	46.4	21	2	US-08-783-288-38	Sequence 38, App
c 38	13	46.4	21	4	US-08-890-640-38	Sequence 38, App
c 39	13	46.4	21	7	5194592-69	Sequence 38, App
c 40	13	46.4	30	6	PCT-US93-01598-27	Sequence 27, App
c 41	13	46.4	48	1	US-08-116-389-16	Sequence 16, App
c 42	13	46.4	48	2	US-08-708-431-16	Sequence 16, App
c 43	13	46.4	48	3	US-08-880-830-16	Sequence 16, App
c 44	13	46.4	48	6	PCT-US94-13895-16	Sequence 16, App
c 45	13	46.4	56	6	PCT-US91-05177-18	Sequence 18, App

ALIGNMENTS

RESULT 1
US-08-361-920-32/C
Sequence 32, Application US/08361920
Patent No. 5457046
GENERAL INFORMATION:
APPLICANT: Woeldike, Helle F.
APPLICANT: Hagen, Frederick
APPLICANT: Hjort, Carsten M.
APPLICANT: Sven, Hastup
TITLE OF INVENTION: An Enzyme Capable of Degrading Cellulose
TITLE OF INVENTION: Or Hemicellulose
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 54570460 No. 5457046disk of No. 5457046th America, Inc.
STREET: 405 Lexington Avenue, 62nd Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/361,920
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/940,860
FILING DATE: 28-OCT-1992
APPLICATION NUMBER: DK 1158/90
FILING DATE: 09-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00124
FILING DATE: 08-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3435, 204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-867-0298
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)


```
RESULT 4
US-08-459-135A-2
; Sequence 2, Application US/08459135A
; Patent No. 595078
; GENERAL INFORMATION:
; APPLICANT: BURHAM, Martin Karl Russel
; APPLICANT: CHOPEA, Ian
; APPLICANT: CRITCHLEY, Ian Alfred
; APPLICANT: KNOWLES, David Justin Charles
; TITLE OF INVENTION: FIBROBLAST BINDING PROTEIN: MONOCLONAL ANTIBODY AND THEIR US
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fastseq, Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,135A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB/94/00215
; FILING DATE: 04-Feb-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmil, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P30591C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: GENOMIC DNA
US-08-459-135A-2

Query Match 53.6%; Score 15; DB 4; Length 34;
Best Local Similarity 78.3%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 cccatgttggtgccaaattg 23
Db 7 CCTTACGTTGTCACGATTGG 29

RESULT 5
US-08-198-431-31/C
; Sequence 31, Application US/08198431
; Patent No. 5605793
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.
; TITLE OF INVENTION: Methods for In Vitro Recombination
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313
```

```
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/198,431
; FILING DATE: 17-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mooi, Leslie
; REGISTRATION NUMBER: 37,047
; REFERENCE/DOCKET NUMBER: 000324-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-8275
; TELEFAX: 415-854-7400
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (oligonucleotide)
US-08-198-431-31

Query Match 50.7%; Score 14.2; DB 1; Length 72;
Best Local Similarity 70.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 ccatgttggtgccaaattg 28
Db 65 CTCTGTTGTACCAATATATGTACG 39

RESULT 6
US-08-564-955-31/C
; Sequence 31, Application US/08564955
; Patent No. 5811238
; GENERAL INFORMATION:
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: CRAMER, ANDREAS M.
; TITLE OF INVENTION: METHODS FOR GENERATING POLYNUCLEOTIDES
; TITLE OF INVENTION: HAVING DESIRED CHARACTERISTICS BY ITERATIVE SELECTION AND
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW
; STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,955
; FILING DATE: 30-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/198,431
; FILING DATE: 17-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/537,874
; FILING DATE: 30-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02126
; FILING DATE: 17-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DUNN, TRACY J.
```



```

: APPLICANT: Nienhuis, Arthur
: APPLICANT: Tolstoshev, Paul
: TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
: TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
: TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
: STREET: 6 Becker Farm Road
: CITY: Roseland
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07068
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch diskette
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: DNA.V2
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/583,276
: FILING DATE: 05-JAN-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/332,444
: FILING DATE: 31-OCT-1994
: APPLICATION NUMBER: 07/887,712
: FILING DATE: 22-MAY-1992
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 35 bases
: TYPE: nucleic acid
: STRANDEDNESS: singular
: TOPOLOGY: linear
: MOLECULE TYPE:
: DESCRIPTION: DNA primer
:
: US-08-583-276-9
:
: Query Match          49.3%; Score 13.8; DB 3; Length 35;
: Best Local Similarity 88.2%; Pred. No. 4.5e+02;
: Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
:
: QY 6 ttgtgtgcacaattg 22
:   | | | | | | | | | |
: Db 19 TATTGTGACAAATTG 35
:
: RESULT 10
: US-08-583-276-10/c
: Sequence 10, Application US/08583276
: Patent No. 5837536
: GENERAL INFORMATION:
: APPLICANT: McDonagh, Kevin T.
: APPLICANT: Nienhuis, Arthur
: APPLICANT: Tolstoshev, Paul
: TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
: TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
: TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
: STREET: 6 Becker Farm Road
: CITY: Roseland
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07068
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch diskette
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: DNA.V2

```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/583,276
: FILING DATE: 05-JAN-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/332,444
: FILING DATE: 31-OCT-1994
: APPLICATION NUMBER: 07/887,712
: FILING DATE: 22-MAY-1992
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 35 bases
: TYPE: nucleic acid
: STRANDEDNESS: singular
: TOPOLOGY: linear
: MOLECULE TYPE:
: DESCRIPTION: DNA primer
:
: US-08-583-276-10
:
: Query Match          49.3%; Score 13.8; DB 3; Length 35;
: Best Local Similarity 88.2%; Pred. No. 4.5e+02;
: Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
:
: QY 6 ttgtgtgcacaattg 22
:   | | | | | | | | | |
: Db 17 TATTGTGACAAATTG 1
:
: RESULT 11
: US-08-743-637B-210/c
: Sequence 210, Application US/08743637B
: Patent No. 5994066
: GENERAL INFORMATION:
: APPLICANT: BERGERON, Michel G.
: APPLICANT: PICARD, Francois J.
: APPLICANT: OUELLETTE, Marc
: APPLICANT: ROY, Paul H.
: TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
: TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
: TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
: TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...
: NUMBER OF SEQUENCES: 273
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: QUARLES & BRADY
: STREET: 411 EAST WISCONSIN AVENUE
: CITY: MILWAUKEE
: STATE: WISCONSIN
: COUNTRY: USA
: ZIP: 53202-4497
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/743,637B
: FILING DATE: 04-NOV-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/526,840
: FILING DATE: 11-SEP-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: BAKER, Jean C.
: REGISTRATION NUMBER: 35,433
: REFERENCE/DOCKET NUMBER: 850586,90012
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (414) 277-5000
: TELEFAX: (414) 277-5591
: INFORMATION FOR SEQ ID NO: 210:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 21 base pairs
: TYPE: nucleic acid

```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-743-637B-210

Query Match          48.6%; Score 13.6; DB 4; Length 21;
Best Local Similarity 80.0%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      7  gtgtgtgcacaattgttca 26
        |||||  |||||  |||||
Db      20  GTGCTGATTAAGGCTCA 1

RESULT 12
US-08-538-875-26
; Sequence 26, Application US/08538875
; Patent No. 5773582
; GENERAL INFORMATION:
; APPLICANT: Shin, Hang-Cheol
; APPLICANT: Shin, Nam-Kyu
; APPLICANT: Lee, Inkyung
; APPLICANT: Kang, Sungzong
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR MOTEINS
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shin, Hang-Cheol
; STREET: Jukong Gochung Apt. 1014-806, Haan-dong
; CITY: Kwangmyung-shi
; STATE: Kyungki-do
; COUNTRY: Republic of Korea
; ZIP: 423-060
; ADDRESSEE: Shin, Nam-Kyu
; STREET: #181-404 Sadang-4-dong, Dongjak-ku
; CITY: Seoul
; STATE:
; COUNTRY: Republic of Korea
; ZIP: 156-094
; ADDRESSEE: Lee, Inkyung
; STREET: 11/2, #302-39 Juan-4-dong, Nam-ku
; CITY: Incheon
; STATE:
; COUNTRY: Republic of Korea
; ZIP: 402-204
; ADDRESSEE: Kang, Sungzong
; STREET: #84-4 Daeshin-dong, Seodaemun-ku
; CITY: Seoul
; STATE:
; COUNTRY: Republic of Korea
; ZIP: 120-160
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5inch 2.0MB storage
; COMPUTER: IBM PC/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/538,875
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,336
; FILING DATE:
; APPLICATION NUMBER: KR 93-1751
; FILING DATE: 9-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX:
; INFORMATION FOR SEQ ID NO: 26:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: primer DNA
US-08-538-875-26

Query Match          48.6%; Score 13.6; DB 2; Length 26;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      6  tgtgtgtgcacaattgttc 25
        ||  |||||  ||  ||
Db      1  TGGTGTGTCACAAAAGGCG 20

RESULT 13
US-09-289-376-21/c
; Sequence 21, Application US/09289376
; Patent No. 6013788
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMAD3 EXPRESSION
; FILE REFERENCE: RTS-0043
; CURRENT APPLICATION NUMBER: US/09/289,376
; CURRENT FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 21
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-289-376-21

Query Match          47.9%; Score 13.4; DB 5; Length 18;
Best Local Similarity 93.3%; Pred. No. 6e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2  cctatgttgtgtcca 16
        |||  |||||  |||
Db      18  CCTGTGTGTGCGCA 4

RESULT 14
US-08-592-406-5
; Sequence 5, Application US/08592406
; Patent No. 5821059
; GENERAL INFORMATION:
; APPLICANT: MINION, F. CHRIS
; APPLICANT: KNOTTSON, Kevin L.
; TITLE OF INVENTION: MYCOPLASMA EXPRESSION SYSTEM
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,406
; FILING DATE: 06-FEB-1996
; CLASSIFICATION: 435
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US93/07407
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 76645/132
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-592-406-5

Query Match 47.9%; Score 13.4; DB 2; Length 44;
Best Local Similarity 93.3%; Pred. No. 7.2e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 ccaaatgtgccaag 28
DB 19 CCCAAATGTGCTCAAG 33

RESULT 15
US-07-635-561A-3
Sequence 3, Application US/07635561A
Patent No. 5244805
GENERAL INFORMATION:
APPLICANT: Miller, Lois K.
TITLE OF INVENTION: Improved Baculovirus Expression Vectors
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner
STREET: 5370 Manhattan Circle, Suite 201
City: Boulder
STATE: CO
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/635,561A
FILING DATE: 19910117
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/353,847
FILING DATE: 17-MAY-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 90/02814
FILING DATE: 17-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: Fether, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 4-90A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
TELEX: 823189
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
US-07-635-561A-3

Query Match 47.9%; Score 13.4; DB 1; Length 97;
Best Local Similarity 73.9%; Pred. No. 8.4e+02;
Matches 17; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 tatgttgggtgccaatgtgtca 26
DB 3 TATCATGAGAGCCAGCTGTGTGA 25

Search completed: June 4, 2000, 16:09:41
Job time: 28060 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:32 ; Search time 1236.38 Seconds

(without alignments)
-22.031 Million cell updates/sec

Title: US-09-164-714-9

Perfect score: 28
Sequence: 1 agatgcacgaacatcaacgtaagaac 28

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*

45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	57.1	79	5 107076	107076 Sequence 2
2	16	57.1	79	5 183771	183771 Sequence 8
3	15.8	56.4	54	5 A57944	A57944 Sequence 10
4	15.8	56.4	54	5 A57945	A57945 Sequence 11
5	15.8	56.4	54	5 A57946	A57946 Sequence 12
6	15.8	56.4	54	5 A57947	A57947 Sequence 13
7	15.8	56.4	54	5 A57948	A57948 Sequence 14
8	15.8	56.4	54	5 A57949	A57949 Sequence 15
9	15.8	56.4	54	5 A57950	A57950 Sequence 16
10	15.8	56.4	54	5 A57951	A57951 Sequence 17
11	15.8	56.4	54	5 A57952	A57952 Sequence 18
12	15.8	56.4	54	5 A57953	A57953 Sequence 19
13	15.8	56.4	54	5 A57954	A57954 Sequence 20
14	15.8	56.4	54	5 A57955	A57955 Sequence 21
15	15.8	56.4	54	5 A57956	A57956 Sequence 22
16	15.8	56.4	54	5 A57957	A57957 Sequence 23
17	15.8	56.4	54	5 A57958	A57958 Sequence 24
18	15.8	56.4	92	5 E05823	E05823 DNA encodin
19	15	53.6	31	5 I37030	I37030 Sequence 43
20	15	53.6	31	5 I93880	I93880 Sequence 43
21	14.8	52.9	53	5 AR053523	AR053523 Sequence
22	14.8	52.9	54	5 A57955	A57955 Sequence 21
23	14.8	52.9	54	5 A57958	A57958 Sequence 24
24	14.8	52.9	54	5 A57961	A57961 Sequence 27
25	14.6	52.1	43	5 A46553	A46553 Sequence 9
26	14.6	52.1	69	5 AR035210	AR035210 Sequence
27	14.6	52.1	69	5 AR035215	AR035215 Sequence
28	14.2	50.7	47	5 A84021	A84021 Sequence 8
29	14.2	50.7	54	5 A57943	A57943 Sequence 9
30	14.2	50.7	54	5 A57953	A57953 Sequence 19
31	14.2	50.7	74	12 R4P2AM25	U52661 Rattus norv
32	14.2	50.7	79	9 HDGCCXC	M87486 Human chrom
33	14.2	50.7	99	5 AR008442	AR008442 Sequence
34	14.2	50.7	99	5 I87434	I87434 Sequence 69
35	13.8	49.3	23	5 A44315	A44315 Sequence 1
36	13.8	49.3	30	5 A44320	A44320 Sequence 6
37	13.8	49.3	42	5 AR026320	AR026320 Sequence
38	13.8	49.3	98	10 HS292873	292873 Homo sapien
39	13.6	48.6	21	5 I76941	I76941 Sequence 7
40	13.6	48.6	23	5 I37347	I37347 Sequence 36
41	13.6	48.6	23	5 I94197	I94197 Sequence 36
42	13.6	48.6	29	5 I83408	I83408 Sequence 9
43	13.6	48.6	32	5 AR035783	AR035783 Sequence
44	13.6	48.6	35	5 A79354	A79354 Sequence 3
45	13.6	48.6	36	5 AR001579	AR001579 Sequence

ALIGNMENTS

RESULT 1
LOCUS 107076 79 bp PAT 02-DEC-1994
DEFINITION Sequence 2 from Patent EP 03144415.
ACCESSION 107076
VERSION 107076.1 GI:590362
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 79)
AUTHORS Namen,A.E., Goodwin,R.G., Lupton,S.D., Mochizuki,D.Y., Price,V.L.
and Deeley,M.C.
TITLE Interleukin-7
JOURNAL Patent: EP 0314415-A2 2 03-MAY-1989;
FEATURES Location/Qualifiers
source 1..79
BASE COUNT 27 a 18 c 22 g 12 t
ORIGIN

Query Match 57.1%; Score 16; DB 5; Length 79;
Best Local Similarity 79.2%; Pred. No. 5.8e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 gatgccaaagcaatcaacgtaag 25
Db 34 GATGACAAAGCAATCGAAGGTAGG 57

RESULT 2
LOCUS 183771 79 bp DNA PAT 10-AUG-1998
DEFINITION Sequence 8 from patent US 5714585.
ACCESSION 183771
VERSION 183771.1 GI:3407301
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 79)
AUTHORS Namen,A.E., Goodwin,R.G., Lupton,S.D. and Mochizuki,D.Y.
TITLE Antibodies that are immunoreactive with Interleukin-7
JOURNAL Patent: US 5714585-A 8 03-FEB-1998;
FEATURES Location/Qualifiers
source 1..79
BASE COUNT 27 a 18 c 22 g 12 t
ORIGIN

Query Match 57.1%; Score 16; DB 5; Length 79;
Best Local Similarity 79.2%; Pred. No. 5.8e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 gatgccaaagcaatcaacgtaag 25
Db 34 GATGACAAAGCAATCGAAGGTAGG 57

RESULT 3
LOCUS A57944 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 10 from Patent EP0743364.
ACCESSION A57944
VERSION A57944.1 GI:3713714
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)

AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 10 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
Other publication FR 2734281 961122.
FEATURES Location/Qualifiers
source 1..54
BASE COUNT 31 a 4 c 13 g 6 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacgtaagaa 27
Db 21 AGAGGAAGACGCAAAACAAAGTAAGAA 47

RESULT 4
LOCUS A57945 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 11 from Patent EP0743364.
ACCESSION A57945
VERSION A57945.1 GI:3713715
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 11 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
Other publication FR 2734281 961122.
FEATURES Location/Qualifiers
source 1..54
BASE COUNT 30 a 4 c 15 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacgtaagaa 27
Db 21 AGAGGAAGACGCAAAACAAAGTAAGAA 47

RESULT 5
LOCUS A57946 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 12 from Patent EP0743364.
ACCESSION A57946
VERSION A57946.1 GI:3713716
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of

JOURNAL HIV-1 mother-foetal transmission
Patent: EP 0743364-A 12 20-NOV-1996;
COMMISSARIAT ENERGIE ATOMIQUE (FR)
COMMENT Other publication FR 2734281 961122.
FEATURES Location/Qualifiers
source 1. .54
/db_xref="taxon:32644"
BASE COUNT 31 a 4 c 14 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGAGCAAAACAAAGTAAGAA 47

RESULT 6
LOCUS A57947 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 13 from Patent EP0743364.
ACCESSION A57947
VERSION A57947.1 GI:3713717
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding
fragments and their application as reactives for risk evaluation of
HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 13 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source 1. .54
/db_xref="taxon:32644"
/db_xref="taxon:32644"
BASE COUNT 31 a 4 c 13 g 6 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGAGCAAAACAAAGTAAGAA 47

RESULT 7
LOCUS A57948 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 14 from Patent EP0743364.
ACCESSION A57948
VERSION A57948.1 GI:3713718
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding
fragments and their application as reactives for risk evaluation of
HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 14 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)

COMMENT Other publication FR 2734281 961122.
FEATURES Location/Qualifiers
source 1. .54
/db_xref="taxon:32644"
BASE COUNT 32 a 5 c 12 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGAGCAAAACAAAGTAAGAA 47

RESULT 8
LOCUS A57949 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 15 from Patent EP0743364.
ACCESSION A57949
VERSION A57949.1 GI:3713719
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding
fragments and their application as reactives for risk evaluation of
HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 15 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source 1. .54
/db_xref="taxon:32644"
/db_xref="taxon:32644"
BASE COUNT 30 a 5 c 14 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGAGCAAAACAAAGTAAGAA 47

RESULT 9
LOCUS A57950 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 16 from Patent EP0743364.
ACCESSION A57950
VERSION A57950.1 GI:3713720
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding
fragments and their application as reactives for risk evaluation of
HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 16 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source 1. .54
Location/Qualifiers

ORIGIN

BASE COUNT 30 a 5 c 14 g 5 t

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacgtaagaa 27
Db 21 AGAGGAAGACAAACAAAGTAAGAA 47

RESULT 10
A57951 54 bp DNA PAT 05-MAR-1998
LOCUS Sequence 17 from Patent EP0743364.
DEFINITION A57951
ACCESSION A57951.1 GI:3713721
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

1 (bases 1 to 54)
Narwa,R. and Roques,P.
Nucleic acid fragments derived from the HIV-1 genome, corresponding of fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
Patent: EP 0743364-A 17 20-NOV-1996;
COMMISSARIAT ENERGIE ATOMIQUE (FR)
Other publication FR 2734281 961122.
Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 29 a 5 c 15 g 4 t 1 others

ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacgtaagaa 27
Db 21 AGAGGAAGACAAACAAAGTAAGAA 47

RESULT 12
A57954 54 bp DNA PAT 05-MAR-1998
LOCUS Sequence 20 from Patent EP0743364.
DEFINITION A57954
ACCESSION A57954.1 GI:3713724
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

1 (bases 1 to 54)
Narwa,R. and Roques,P.
Nucleic acid fragments derived from the HIV-1 genome, corresponding of fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
Patent: EP 0743364-A 20 20-NOV-1996;
COMMISSARIAT ENERGIE ATOMIQUE (FR)
Other publication FR 2734281 961122.
Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 30 a 6 c 14 g 4 t

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGAGCAACAAAGTAAGAA 47

RESULT 14
A57957
LOCUS A57957 54 bp DNA PAT 05-MAR-1998

DEFINITION Sequence 23 from Patent EP0743364.
ACCESSION A57957
VERSION A57957.1 GI:3713727

KEYWORDS
SOURCE
ORGANISM
REFERENCE

1 (bases 1 to 54)
Narwa,R. and Roques,P.
Nucleic acid fragments derived from the HIV-1 genome, corresponding
fragments and their application as reactives for risk evaluation of
HIV-1 mother-foetal transmission

Patent: EP 0743364-A 23 20-NOV-1996;
COMMISSARIAT ENERGIE ATOMIQUE (FR)
Other publication FR 2734281 961122.

COMMENT
FEATURES
Source
Location/Qualifiers

1. .54

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 30 a 6 c 13 g 5 t

ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGAGCAACAAAGTAAGAA 47

RESULT 15
A57959
LOCUS A57959 54 bp DNA PAT 05-MAR-1998

DEFINITION Sequence 25 from Patent EP0743364.
ACCESSION A57959
VERSION A57959.1 GI:3713729

KEYWORDS
SOURCE
ORGANISM
REFERENCE

1 (bases 1 to 54)
Narwa,R. and Roques,P.
Nucleic acid fragments derived from the HIV-1 genome, corresponding
fragments and their application as reactives for risk evaluation of
HIV-1 mother-foetal transmission

Patent: EP 0743364-A 25 20-NOV-1996;
COMMISSARIAT ENERGIE ATOMIQUE (FR)
Other publication FR 2734281 961122.

COMMENT
FEATURES
Source
Location/Qualifiers

1. .54

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 30 a 4 c 15 g 5 t

ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGAGCAACAAAGTAAGAA 47

Search completed: June 4, 2000, 16:05:33
Job time: 2785 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:06 ; Search time 322.35 Seconds
(without alignments)
21.732 Million cell updates/sec

Title: US-09-164-714-9

Perfect score: 28

Sequence: 1 agatgccaaagcaatcaacgtaagaac 28

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.8	56.4	54	1 T43658	HIV-1 matrix prote
2	15.8	56.4	54	1 T43659	HIV-1 matrix prote
3	15.8	56.4	54	1 T43662	HIV-1 matrix prote
4	15.8	56.4	54	1 T43643	HIV-1 matrix prote
5	15.8	56.4	54	1 T43644	HIV-1 matrix prote
6	15.8	56.4	54	1 T43645	HIV-1 matrix prote
7	15.8	56.4	54	1 T43646	HIV-1 matrix prote
8	15.8	56.4	54	1 T43647	HIV-1 matrix prote
9	15.8	56.4	54	1 T43648	HIV-1 matrix prote
10	15.8	56.4	54	1 T43649	HIV-1 matrix prote
11	15.8	56.4	54	1 T43650	HIV-1 matrix prote
12	15.8	56.4	54	1 T43652	HIV-1 matrix prote
13	15.8	56.4	54	1 T43654	HIV-1 matrix prote
14	15.8	56.4	54	1 T43655	HIV-1 matrix prote
15	15.8	56.4	54	1 T43657	HIV-1 matrix prote
16	15.8	56.4	92	1 Q53311	Hybrid antigen pro
17	15.2	54.3	38	1 N94509	Probe for N-termin
18	15.2	54.3	38	1 V83523	PCR primer used to
19	15.2	54.3	38	1 V83525	PCR primer used to
20	15.2	54.3	97	1 T19416	Human gene signatu
21	15.2	53.6	31	1 Q93504	Human stromelysin
22	14.8	52.9	47	1 X52584	Human genome biall
23	14.8	52.9	47	1 X52596	Human genome biall
24	14.8	52.9	47	1 X52546	Human genome biall
25	14.8	52.9	47	1 X52546	Human genome biall
26	14.8	52.9	53	1 V68220	PCR primer used to
27	14.8	52.9	54	1 Q86781	SRF-1 mimetic prim
28	14.8	52.9	54	1 Q74004	SRF-1 mimetic PCR
29	14.8	52.9	54	1 T43653	HIV-1 matrix prote
30	14.8	52.9	54	1 T43653	HIV-1 matrix prote
31	14.6	52.1	54	1 T43656	HIV-1 matrix prote
32	14.6	52.1	69	1 T01299	PCR primer oligo D
33	14.6	52.1	69	1 Q30897	Primer 312-69. New
34	14.6	52.1	69	1 Q30892	Primer 312-64. New

C 35	14.6	52.1	100	1 T42735	Primer J089 for hu
C 36	14.4	51.4	60	1 T62581	PCR primer for cys
C 37	14.2	50.7	47	1 V68272	Penicillium chryso
C 38	14.2	50.7	54	1 T43651	HIV-1 matrix prote
C 39	14.2	50.7	54	1 T43641	HIV-1 matrix prote
C 40	14.2	50.7	99	1 V10203	Stealth virus nucl
C 41	14.2	50.7	99	1 V12016	Stealth virus plas
C 42	14	50.0	99	1 T30883	Primer 13 for 95 k
C 43	13.8	49.3	23	1 Q88336	PCR primer OTG3042
C 44	13.8	49.3	30	1 Q88341	PCR primer OTG5988
C 45	13.8	49.3	37	1 V39939	Streptococcus pneu

ALIGNMENTS

RESULT 1	
T43658	
ID T43658 standard; DNA; 54 BP.	
AC T43658;	
DT 19-AUG-1997 (first entry)	
DE HIV-1 matrix protein p17 gene fragment 2754.	
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;	
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.	
OS Human immunodeficiency virus type 1.	
PN Ep-743364-A2.	
PD 20-NOV-1996.	
PE 17-MAY-1996; 401084.	
PR 18-MAY-1995; FR-005914.	
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.	
PI Narwa R, Roques P;	
DR WPI: 96-507733/51.	
PT Human immunodeficiency virus p17 gene fragments, derived proteins	
PT and antibodies - useful for assessing the risk of maternal	
PT transmission of HIV-1 infection	
PS Claim 3; Page 27; 46pp. French.	
CC This sequence is a specifically claimed example of 21-90 nucleotide	
CC long nucleic acid fragments, derived from the gene encoding part of the	
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at	
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;	
CC Y6 = ATA, TTA, CTG, GTA, CTA, GTC or ATG; Y7 = GAG or GAA; Y8 = GAA or	
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or	
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5	
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,	
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.	
CC The new nucleic acid fragments can be used as reagents for determining	
CC and assessing the risk of maternal-fetal transmission of HIV-1, using	
CC standard hybridisation or immuno assays. The presence of such sequences	
CC in maternal blood is strongly correlated with transmission of infection.	
SO Sequence 54 BP; 30 A; 4 C; 15 G; 5 T;	
Query Match 56.4%; Score 15.8; DB 1; Length 54;	
Best Local Similarity 74.1%; Pred. No. 2.1e+02;	
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
QY 1 agatgccaaagcaatcaacgtaagaac 27	
DB 21 AGAGCAAGAGCAACAAACAAAGTAACAA 47	
RESULT 2	
T43659	
ID T43659 standard; DNA; 54 BP.	
AC T43659;	
DT 19-AUG-1997 (first entry)	
DE HIV-1 matrix protein p17 gene fragment 2826.	
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;	
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.	
OS Human immunodeficiency virus type 1.	
PN Ep-743364-A2.	
PD 20-NOV-1996.	
PE 17-MAY-1996; 401084.	

PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R. Roques P;
DR WPI; 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 27; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 30 A; 4 C; 15 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 agatgccaaagcaaatcaacgtaagaa 27
||| | ||||| ||| ||||| |||
Db 21 AGAGGAGAGCAACAAACAAAGTAAGAA 47

RESULT 3
T43642
ID T43642 standard; DNA; 54 BP.
AC T43642;
DE 19-AUG-1997 (first entry)
DT HIV-1 matrix protein p17 gene fragment ARI.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KM detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R. Roques P;
DR WPI; 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 23; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 31 A; 4 C; 13 G; 6 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 agatgccaaagcaaatcaacgtaagaa 27
||| | ||||| ||| ||||| |||
Db 21 AGAGGAGAGCAACAAACAAAGTAAGAA 47

RESULT 4
T43643
ID T43643 standard; DNA; 54 BP.
AC T43643;
DE 19-AUG-1997 (first entry)
DT HIV-1 matrix protein p17 gene fragment BOI.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KM detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
FH Key Location/Qualifiers
FT mat_peptide 1..54
FT /tag= a
FT /product= BOI_peptide
PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R. Roques P;
DR WPI; 96-507733/51.
DR P-PSDB; W06610.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 23; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 30 A; 4 C; 15 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 agatgccaaagcaaatcaacgtaagaa 27
||| | ||||| ||| ||||| |||
Db 21 AGAGGAGAGCAACAAACAAAGTAAGAA 47

RESULT 5
T43644
ID T43644 standard; DNA; 54 BP.
AC T43644;
DE 19-AUG-1997 (first entry)
DT HIV-1 matrix protein p17 gene fragment DMU.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KM detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R. Roques P;
DR WPI; 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal

PT transmission of HIV-1 infection
PS Claim 3: Page 24; 46pp: French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GTC or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CCA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 31 A; 4 C; 14 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGCAAGACGCAAAACAAAGTAAAGAA 47

RESULT 6
T43645
ID T43645 standard; DNA; 54 BP.
AC T43645;
DE 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment PAL.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection: maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;
DR WPI; 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection.
PS Claim 3: Page 24; 46pp: French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GTC or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CCA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 31 A; 4 C; 13 G; 6 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGCAAGACGCAAAACAAAGTAAAGAA 47

RESULT 7

T43646
ID T43646 standard; DNA; 54 BP.
AC T43646;
DE 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment RYO.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection: maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;
DR WPI; 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection.
PS Claim 3: Page 24; 46pp: French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GTC or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CCA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
SQ Sequence 54 BP; 32 A; 5 C; 12 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGCAAGACGCAAAACAAAGTAAAGAA 47

RESULT 8
T43647
ID T43647 standard; DNA; 54 BP.
AC T43647;
DE 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment FLO.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection: maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;
DR WPI; 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection.
PS Claim 3: Page 24; 46pp: French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GTC or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CCA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.

CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 30 A; 5 C; 14 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| ||||| |||
Db 21 AGAGGAAGACCAAAACAAAGTAAGAA 47

RESULT 9
T43648 19-AUG-1997 (first entry)
AC T43648; standard: DNA; 54 BP.

DE HIV-1 matrix protein p17 gene fragment 4501.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
CC detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
FH Key
FT mat_peptide 1..54
FT Location/Qualifiers

FT /*tag= a
FT /product= 4501_peptide

PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R., Roques P.
DR WPI: 96-507733/51.

PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 24; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 30 A; 5 C; 14 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| ||||| |||
Db 21 AGAGGAAGACCAAAACAAAGTAAGAA 47

RESULT 10
T43649 19-AUG-1997 (first entry)
AC T43649; standard: DNA; 54 BP.

DE HIV-1 matrix protein p17 gene fragment FAI.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;

KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.

PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R., Roques P.
DR WPI: 96-507733/51.

PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 25; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not AAT or AAC.
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 31 A; 4 C; 14 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| ||||| |||
Db 21 AGAGGAAGACCAAAACAAAGTAAGAA 47

RESULT 11
T43650 19-AUG-1997 (first entry)
AC T43650; standard: DNA; 54 BP.

DE HIV-1 matrix protein p17 gene fragment HAR.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
CC detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
FH Key
FT mat_peptide 1..54
FT Location/Qualifiers

FT /*tag= a
FT /product= HAR_peptide

PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.

PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

PI Narwa R., Roques P.

DR WPI: 96-507733/51.

PT Human immunodeficiency virus p17 gene fragments, derived proteins

PT and antibodies - useful for assessing the risk of maternal

PT transmission of HIV-1 infection

PS Claim 3; Page 25; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining

CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP: 29 A; 5 C; 15 G; 4 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAGAGCGCAAAACAAAGTAAGAA 47

RESULT 12
T43652
ID T43652 standard; DNA; 54 BP.
AC T43652;
DT 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment CHET.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.

PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;

DR WPI: 96-507733/51.

PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
transmission of HIV-1 infection
PS Claim 3; Page 26; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, GTG, CTA, GTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP: 30 A; 4 C; 15 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;

Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAGAGCGCAAAACAAAGTAAGAA 47

RESULT 13

T43654
ID T43654 standard; DNA; 54 BP.

AC T43654;

DT 19-AUG-1997 (first entry)

DE HIV-1 matrix protein p17 gene fragment SIW.

KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.

OS Human immunodeficiency virus type 1.

PH Key Location/Qualifiers

FT mat_peptide 1..54

FT /*tag= a

FT /product= SIW_peptide

PN EP-743364-A2.

PD 20-NOV-1996.

PF 17-MAY-1996; 401084.

PR 18-MAY-1995; FR-005914.

PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

PI Narwa R, Roques P;

DR WPI: 96-507733/51.

DR P-PSDB; W06614.

PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal

transmission of HIV-1 infection

PS Claim 3; Page 26; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, GTG, CTA, GTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP: 30 A; 6 C; 14 G; 4 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 agatgccaaagcaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAGAGCGCAAAACAAAGTAAGAA 47

RESULT 14

T43655
ID T43655 standard; DNA; 54 BP.

AC T43655;

DT 19-AUG-1997 (first entry)

DE HIV-1 matrix protein p17 gene fragment MOE.

KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.

OS Human immunodeficiency virus type 1.

PH Key Location/Qualifiers

FT mat_peptide 1..54

FT /*tag= a

FT /product= MOE_peptide

PN EP-743364-A2.

PD 20-NOV-1996.

PF 17-MAY-1996; 401084.

PR 18-MAY-1995; FR-005914.

PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

PI Narwa R, Roques P;

DR WPI: 96-507733/51.

DR P-PSDB; W06615.

PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal

transmission of HIV-1 infection

PS Claim 3; Page 26; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, GTG, CTA, GTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using

CC standard hybridisation or immuno assays. The presence of such sequences
 CC in maternal blood is strongly correlated with transmission of infection.
 SQ Sequence 54 BP; 30 A; 6 C; 13 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;

Best Local Similarity 74.1%; Pred. No. 2.1e+02;

Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 agatgccaaagcaatcaacgtaagaa 27
 ||| | ||||| ||| |||||
 Db 21 AGAGGAAGACCAACCAAGTAAGAA 47

RESULT 15

T43657

T43657 standard; DNA; 54 BP.

AC T43657;

DT 19-AUG-1997 (first entry)

DE HIV-1 matrix protein p17 gene fragment 4541.

KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;

OS Human immunodeficiency virus type 1.

PN EP-743364-A2.

PD 20-NOV-1996.

PF 17-MAY-1996; 401084.

PR 18-MAY-1995; FR-005914.

PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

PI Narwa R. Roques P;

DR MPI; 96-507733/51.

PT Human immunodeficiency virus p17 gene fragments, derived proteins

PT and antibodies - useful for assessing the risk of maternal

PT transmission of HIV-1 infection

PS Claim 3; Page 27; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide

CC long nucleic acid fragments, derived from the gene encoding part of the

CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at

CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;

CC Y6 = ATA, TTA, CTG, GTA, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or

CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or

CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5

CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,

CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.

CC The new nucleic acid fragments can be used as reagents for determining

CC and assessing the risk of maternal-foetal transmission of HIV-1, using

CC standard hybridisation or immuno assays. The presence of such sequences

CC in maternal blood is strongly correlated with transmission of infection.

SQ Sequence 54 BP; 30 A; 4 C; 15 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;

Best Local Similarity 74.1%; Pred. No. 2.1e+02;

Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 agatgccaaagcaatcaacgtaagaa 27
 ||| | ||||| ||| |||||
 Db 21 AGAGGAAGACCAACCAAGTAAGAA 47

Search completed: June 4, 2000, 16:24:08
 Job time: 28904 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:09 ; Search time 4521.53 Seconds
(without alignments)
25.100 Million cell updates/sec

Title: US-09-164-714-9

Perfect score: 28

Sequence: 1 agatgccagcaatcaatcaagtaagaac 28

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026511650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est20:*
53: em_est21:*
54: em_est12:*
55: em_est13:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est13:*
60: gb_est14:*
61: gb_est15:*
62: gb_est16:*
63: gb_est17:*
64: gb_est18:*
65: em_est17:*
66: em_est18:*
67: em_est29:*
68: em_est30:*
69: gb_est19:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est132:*
77: em_est133:*
78: em_est134:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss11:*
83: gb_gss12:*
84: gb_gss13:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result

*
Query

No.	Score	Match	Length	DB	ID	Description
1	15.2	54.3	62	38	AA746404	AA746404.nw62401.s
2	15.2	54.3	62	38	AA746400	AA746400.nw62c07.s
3	15.2	54.3	72	30	AA744996	AA744996.nw29407.s
4	15.2	54.3	75	38	AA746399	AA746399.nw62c05.s
5	15.2	54.3	75	38	AA746402	AA746402.nw62c11.s
6	15.2	54.3	83	37	AA681983	AA681983.vu76c07.r
7	15.2	54.3	89	20	D11978	D11978.HUMOS12E03
8	15.2	54.3	94	20	D12000	D12000.HUMOS13A08
9	15.2	53.6	84	83	AF088162	AF088162.Homo.sapi
10	14.8	52.9	81	40	AA930201	AA930201.vst1e03.r
11	14.6	52.1	90	47	AI536307	AI536307.col14a09.x
12	14.4	51.4	60	29	AA161708	AA161708.MBAFC9F0
13	14.4	51.4	82	64	AA057193	AA057193.ca02a05.y
14	14.4	51.4	87	40	AA916294	AA916294.on22b10.s
15	14.4	51.4	94	60	AI784864	AI784864.SWAMCAG30
16	14.4	51.4	99	42	AI082999	AI082999.SWAMCAG18
17	14.2	50.7	43	37	AA682184	AA682184.hn51n12.s
18	14.2	50.7	85	40	AA910924	AA910924.ok85e02.s
19	14.2	50.7	92	24	H99072	H99072.yv90h05.s1
20	14.2	50.0	52	48	AI584790	AI584790.fB83408.y
21	14.2	50.0	73	33	AA398108	AA398108.zt58d08.s
22	14.2	50.0	82	62	AI906839	AI906839.rc-br126-
23	13.8	49.3	29	20	TI7517	TI7517.gsr.m46.Tne
24	13.8	49.3	80	30	AA249355	AA249355.j1284.seq
25	13.8	49.3	85	30	AA231275	AA231275.mw32c03.r
26	13.8	49.3	100	23	H19067	H19067.yn51b08.r1
27	13.6	48.6	64	81	AA049557	AA049557.SMOV3CMAM
28	13.6	48.6	67	44	AI286661	AI286661.uB95402.r
29	13.6	48.6	76	35	C58767	C58767.C58767.YuJ1
30	13.6	48.6	76	62	AI915855	AI915855.wg95a07.x
31	13.6	48.6	78	30	AA230217	AA230217.nc13f02.r
32	13.6	48.6	88	63	AI960312	AI960312.sc81g04.y
33	13.6	48.6	91	79	AA286463	AA286463.LG1.333.B
34	13.6	48.6	92	28	C20817	C20817.HUMGS004S6
35	13.6	48.6	94	31	AA291480	AA291480.zt40b04.s
36	13.6	48.6	94	45	AI365056	AI365056.q112b12.x
37	13.6	48.6	97	103	AA0248664	AA0248664.F13111.SP
38	13.6	48.6	99	24	H74881	H74881.226.Deltoid
39	13.6	48.6	100	34	AA509111	AA509111.MBAFCX7D0
40	13.4	47.9	46	94	AA073659	AA073659.BP(2)2525
41	13.4	47.9	55	69	AA012242	AA012242.s085b11.y
42	13.4	47.9	58	41	AU008799	AU008799.AU008799
43	13.4	47.9	59	63	AI988246	AI988246.sc98a02.y
44	13.4	47.9	64	37	AA715443	AA715443.nv53b08.r
45	13.4	47.9	67	39	AA905978	AA905978.oj89c03.s

ALIGNMENTS

```

RESULT 1
AA746404 62 bp mRNA 27-JAN-1998
DEFINITION nw62601.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1251169 3',
ACCESSION AA746404
VERSION AA746404.1 GI:2786390
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 62)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE Tumor Gene Index
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,

```

Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bdrp/image/image.html

Insert length: 471 Std Error: 0.00
 Seq primer: -40m13 fwd. RT from Amersham
 High quality sequence stop: 57.
 Location/Qualifiers
 1. 62
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1251169"
 /clone_1lb="NCI_CGAP_GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+, IgD-),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 [5'-TGTTCACATTCGAGTGGGACGGCCGTCATTTTTTTTTTTTTT-
 3']. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 25 a 11 c 17 g 9 t

Query Match 54.3%; Score 15.2; DB 38; Length 62;

Best Local Similarity 71.4%; Pred. No. 7.3e+03;

Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 agatgccagcaatcaacgtaagaac 28
 ||||| || ||||| ||||| ||
 Db 15 AGATGCAGAGAACTCAAGCTAAATAC 42

```

RESULT 2
AA746400 63 bp mRNA EST 27-JAN-1998
DEFINITION nw62c07.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1251180 3',
ACCESSION AA746400
VERSION AA746400.1 GI:2786386
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 63)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE Tumor Gene Index
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html

Insert Length: 648 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 63.

FEATURES

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1251180"
/clone_lib="NCI-CGAP-GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, ID-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CEBR). cDNA synthesis was
primed with a Not I - cll9g(dT) primer
[5'-TGTTCCAACTCGACGTGGGCGGCCGCTCAATTTTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

```

Query Match	54.38;	Score 15.2;	DB 38;	Length 63;
Best Local Similarity	71.4%;	Pred. No. 7.3e+03;		
Matches 20; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0

RESULT	3						
AA244996							
LOCUS		72 bp	mRNA	EST		12-MAR-1997	
DEFINITION		mv29g07.r1 Guaymotord Beier mouse kidney day 0					mus musculus cDNA
		clone IMAGE:656461 5',	mRNA sequence.				
ACCESSION	AA244996						
VERSION	AA244996.1	GI:1875731					
KEYWORDS	EST.						
SOURCE	house mouse.						

FEATURES
source

	FEATURES	SOURCE
	Location/Qualifiers	
1..72	/organism="Mus musculus"	
	/strain="C57Bl/6J"	
	/db_xref="taxon:10090"	
	/clone_image="656461"	
	/clone_lib="Guaywoodford Beier mouse kidney day 0"	
	/clone_type="Kidney"	
	/dev_stage="newborn (day 0)"	
	/lab_host="SOLR (kanamycin resistant)"	
	/note="Organ: kidney; Vector: pluscript SK-; Site:1; Ecorr: Site:2; XhoI: Cloned unidirectionally. Primer: Oligo df. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAAATCGCACGAC 3' ~3' adaptor sequence: 5' CTCGAGCTTTT TTTTTTTTTT 3' Library provided Lisa Guay-Woodford."	
BASE COUNT	33 a	13 t
ORIGIN	11 c	14 g 1 others

Query Match	54.3%	Score 15.2;	DB 30;	Length 72;
Best Local Similarity	71.4%	Pred. No. 7.4e+03;		
Matches 20; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;
OY	1	agatgcgaagcaatcaacggtgaagac	28	
DB	45	agatgaagactaatataaagtaagaac	72	

RESULT		4				
AA746399						
LOCUS						
DEFINITION	n62cc05.s1 NCI_CGAP_GCB1 Homo sapiens CDNA clone IMAGE:1251176 3'					
ACCESSION	mRNA sequence.					
VERSION	AA746389					
SOURCE	AA746399.1 GI:2786385 EST.					
ORGANISM	human. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.					
REFERENCE	NCl-CGAP http://www.ncbi.nlm.nih.gov/ncigap .					
AUTHORS	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index					
TITLE	Unpublished (1997)					
JOURNAL	Contact: Robert Strausberg, Ph.D.					
COMMENT						

```

Insert Length: 363   Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 63.
location/Qualifiers
    source
        1. 75

```

TITLE	The Washu-HHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	On Dec 30, 1996 this sequence version replaced gi:1530953

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through ILMN; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:402309
Seq primer: -28m3 rev1 ET from Amersham
High quality sequence stop: 57.

germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-GTGTACCACTCTGAAGTGGAGCGCGCTCATTTTCTTTTCTTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRTV3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 29 a 14 c 20 g 12 t

ORIGIN

Query Match 54.3%; Score 15.2; DB 38; Length 75;
Best Local Similarity 71.4%; Pred. No. 7.4e+03;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 agatgccagcaaatcaacggtgaagaac 28
||||| || ||||| ||||| ||
Db 15 AGATGCAGAGAACTCAAGCTAATATC 42

RESULT 5
AA746402 75 bp mRNA EST 27-JAN-1998
LOCUS nw62c11.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1251188 3',
DEFINITION mRNA sequence.
ACCESSION AA746402
VERSION AA746402.1 GI:2786388
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 75)
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Cloned distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnl.gov/db/rlp/image/image.html

Insert Length: 664 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 63.
location/Qualifiers
1. 75
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1251188"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pRTV3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-GTGTACCACTCTGAAGTGGAGCGCGCTCATTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pRTV3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 29 a 14 c 20 g 12 t

ORIGIN

Query Match 54.3%; Score 15.2; DB 38; Length 75;
Best Local Similarity 71.4%; Pred. No. 7.4e+03;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 agatgccagcaaatcaacggtgaagaac 28
||||| || ||||| ||||| ||
Db 15 AGATGCAGAGAACTCAAGCTAATATC 42

RESULT 6
AA681983 83 bp mRNA EST 05-DEC-1997
LOCUS vu76c07.r1 StrataGene mouse skin (#937313) Mus musculus cDNA clone
DEFINITION IMAGE:1197324 5', mRNA sequence.
ACCESSION AA681983
VERSION AA681983.1 GI:2664123
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 83)
JOURNAL Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Thelning,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
On Sep 12, 1996 this sequence version replaced gi:1392917.
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LINL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:644420
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 76.
location/Qualifiers
1. 83
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone_image="1197324"
/clone_lib="Stratagene mouse skin (#937313)"
/sex="females"
/tissue_type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: skin; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GATTTCGGCAGCAG 3' -3' adaptor
sequence: 5' CTCGAGCTTTTCTTTTCTTTTCTTTT 3"

FEATURES
source

BASE COUNT 38 a 12 c 14 g 19 t

ORIGIN

Query Match 54.3%; Score 15.2; DB 37; Length 83;
Best Local Similarity 71.4%; Pred. No. 7.5e+03;

Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 agatgccaaagaatcaacggtgaagac 28
 ||||| | ||||| | ||||| |
 Db 22 AGATGAAAGACTAATTAAGTAAGAAC 49

RESULT 7
 D11978 89 bp mRNA EST 02-DEC-1992
 LOCUS HUM0512E03 Liver Hepg2 cell line. Homo sapiens cDNA clone s12e03,
 DEFINITION mRNA sequence.
 D11978
 ACCESSION D11978.1 GI:2148809
 VERSION D11978.1 GI:2148809
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 89)
 Okubo,K., Hori,N., Matoba,R., Niyama,T., Fukushima,A., Kojima,Y.
 and Matsubara,K.
 Large scale cDNA sequencing for analysis of quantitative and
 qualitative aspects of gene expression
 Nature Genet. 2, 173-179 (1992)

JOURNAL 94258199
 MEDLINE
 COMMENT Contact: Kousaku Okubo, Naohiro Hori, Ryo Matoba, Toshiyuki
 Niyama, Atsushi Fukushima, Yuko Kojima & Kenichi Matsubara
 Institute for Molecular and Cellular Biology
 Osaka University
 1-3 Yamada-oka, Suita, Osaka 565, Japan.
 Location/Qualifiers
 1..89
 /organism="Homo sapiens"
 /db_xref="GDB:D058180E"
 /db_xref="taxon:9606"
 /clone="s12e03"
 /clone_lib="Liver Hepg2 cell line."
 /lab_host="E.coli"
 /note="3'-directed regional cDNA library. Cleaved by MboI
 and transformed into E.coli."

BASE COUNT 14 a 21 c 22 g 32 t
 ORIGIN

Query Match 54.3%; Score 15.2; DB 20; Length 89;
 Best Local Similarity 71.4%; Pred. No. 7.5e+03;
 Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 agatgccaaagaatcaacggtgaagac 28
 ||||| | ||||| | ||||| |
 Db 73 AAATGCCAGCAATCAACAGGAGGAC 46

RESULT 8
 D12000/c 94 bp mRNA EST 02-DEC-1992
 LOCUS HUM0513A08 Liver Hepg2 cell line. Homo sapiens cDNA clone s13a08,
 DEFINITION mRNA sequence.
 D12000
 ACCESSION D12000.1 GI:2148823
 VERSION D12000.1 GI:2148823
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 94)
 Okubo,K., Hori,N., Matoba,R., Niyama,T., Fukushima,A., Kojima,Y.
 and Matsubara,K.
 Large scale cDNA sequencing for analysis of quantitative and
 qualitative aspects of gene expression
 Nature Genet. 2, 173-179 (1992)

JOURNAL 94258199
 MEDLINE

COMMENT Contact: Kousaku Okubo, Naohiro Hori, Ryo Matoba, Toshiyuki
 Niyama, Atsushi Fukushima, Yuko Kojima & Kenichi Matsubara
 Institute for Molecular and Cellular Biology
 Osaka University
 1-3 Yamada-oka, Suita, Osaka 565, Japan.
 Location/Qualifiers
 1..94
 /organism="Homo sapiens"
 /db_xref="GDB:D058194E"
 /db_xref="taxon:9606"
 /clone="s13a08"
 /clone_lib="Liver Hepg2 cell line."
 /lab_host="E.coli"
 /note="3'-directed regional cDNA library. Cleaved by MboI
 and transformed into E.coli."

BASE COUNT 18 a 21 c 33 t 1 others
 ORIGIN

Query Match 54.3%; Score 15.2; DB 20; Length 94;
 Best Local Similarity 71.4%; Pred. No. 7.5e+03;
 Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 agatgccaaagaatcaacggtgaagac 28
 ||||| | ||||| | ||||| |
 Db 73 AAATGCCAGCAATCAACAGGAGGAC 46

RESULT 9
 AF088162/c 84 bp DNA GSS 10-DEC-1998
 LOCUS AF088162 Homo sapiens chromosome 11 clone 49D4E3 map 11q13, genomic survey
 DEFINITION sequence.
 AF088162
 ACCESSION AF088162.1 GI:3982872
 VERSION AF088162.1 GI:3982872
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 84)
 Gerhard,D.S. and Zhu,S.
 A transcript map of an 800 kb region on human chromosome 11q13,
 part of the candidate region for SCAs and BBS1
 Hum. Genet. (1999) In press

JOURNAL 2 (bases 1 to 84)
 AUTHORS Gerhard,D.S. and Zhu,S.
 TITLE Direct Submission
 JOURNAL Submitted (31-AUG-1998) Department of Genetics, Washington
 University School of Medicine, 4566 Scott Avenue, Box 8232, St.
 Louis, MO 63110, USA
 Location/Qualifiers
 1..84
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="11q13"
 /clone="49D4E3"
 /chromosome="11"

BASE COUNT 24 a 17 c 10 g 33 t
 ORIGIN

Query Match 53.6%; Score 15; DB 83; Length 84;
 Best Local Similarity 78.3%; Pred. No. 9.1e+03;
 Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 gccaaagaatcaacggtgaagaa 27
 ||||| | ||||| | ||||| |
 Db 34 GCCAAGCACATCAAAAGAACGAA 12

RESULT 10
 AA930201

LOCUS	AA930201	81 bp	mRNA	EST	23-APR-1998
DEFINITION	vs71e03.t1 Stratagene mouse skin (#937313) Mus musculus cDNA clone IMAGE:1151740 5' similar to gb:U15647_cds1 Mus musculus (MOUSE);/, mRNA sequence.				
ACCESSION	AA930201				
VERSION	AA930201.1	GI:3079794			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.				
TITLE	The Mashu-HMI Mouse EST Project				
JOURNAL	Unpublished (1996)				
COMMENT	On Jan 19, 1998 this sequence version replaced gi:2286367. Contact: Marra M/Mouse EST Project Mashu-HMI Mouse EST Project Washington University School of MedicineP 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (infoimage.lnl.gov) for further information. MGJ:624948 Seq primer: -28ml3 rev1 ER from Amersham.				
FEATURES	Location/Qualifiers				
source	1..81				
	/organism="Mus musculus"				
	/strain="C57BL/6"				
	/db_xref="taxon:10090"				
	/clone="IMAGE:1151740"				
	/clone_1lb="Stratagene mouse skin (#937313)"				
	/sex="females"				
	/tissue_type="whole skin"				
	/dev_stage="11 weeks old"				
	/lab_host="SOAR (kanamycin resistant)"				
	/note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Whole skin from 11 week old C57BL/6 female muce Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCGAGG 3' ~3' adaptor sequence: 5' CTCAGCTTTTTTTTTTTT 3'"				
BASE COUNT	35 a	21 c	9 g	16 t	
ORIGIN					
Query Match	52.9%	Score 14.8;	DB 40;	Length 81;	
Best Local Similarity	73.1%;	Pred. NO.1.le=04;			
Matches 19;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;	
OY	1 agatccgaagcaaatcaccgtaaga 26				
Db	41 AATGCATAATCAACAACCCGTGAGA 66				
RESULT 11					
A1536907					
LOCUS	A1536907	90 bp	mRNA	EST	12-MAY-1999
DEFINITION	tol4ab09.x1 NC1_CGAP_ut2 Homo sapiens cDNA clone IMAGE:2179000 3', mRNA sequence.				
ACCESSION	A1536907				
VERSION	A1536907.1	GI:4451042			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiuidae; Homo.				

REFERENCE	1 (bases 1 to 90)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	On May 18, 1998 this sequence version replaced g1:3138691. Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert_Strausberg@nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: www-bio.llnl.gov/bdrp/image/image.html
FEATURES	Insert Length: 1143 Std Error: 0.00 Seq primer: -40UP from Glbco High quality sequence stop: 85 POLYA-No.
SOURCE	Location/Qualifiers 1..90 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2179000" /clone_1id="NCI-CGAP-UF2" /tissue_type="moderately-differentiated endometrial adenocarcinoma, 3 pooled tumors" /lab_host="DH10B" /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.85 kb. Life Technologies catalog #: 11539-012"
BASE COUNT	34 a 16 c 13 g 27 t
ORIGIN	
Query Match	52.1%; Score 14.6; DB 47; Length 90;
Best Local Similarity	81.0%; Pred. No. 1.4e+04;
Matches 17; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
Ox	7 caagcaatacaagcgttaaga 27
Db	17 CAAGCACAACAACGTAAGAA 37
RESULT 12	
LOCUS	AA161708 60 bp mRNA EST 17-DEC-1996
DEFINITION	MBAFCE9P05T3 Brugia malayi adult female cDNA (S9696MLW-BMAF) Brugia malayi cDNA clone AFCE9F05 5', mRNA sequence.
ACCESSION	AA161708
VERSION	AA161708.1 GI:1737540
KEYWORDS	EST.
SOURCE	Brugia malayi.
ORGANISM	Brugia malayi. Eukaryota; Metazoa; Nematoda; Secernentea; Spirurina; Spirurida; Filarioida; Onchoercidae; Brugia.
REFERENCE	1 (bases 1 to 60)
AUTHORS	Blaxter/M.L., Waterfall,M., Daub,J.J., Lizotte,M., Baron,L. and Jones,S.J.
TITLE	Genes expressed in adult female Brugia malayi
COMMENT	Unpublished (1996) On May 18, 1995 this sequence version replaced g1:811474. Contact: Blaxter ML Institute of Cell, Animal and Population Biology University of Edinburgh Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JF, UK. Tel.: +44 131 650 6760 Fax: +44 131 670 5450

Email: mark.blaxter@ed.ac.uk
The ABI trace of this sequence can be viewed at
<http://www.sanger.ac.uk/brugia/ABC/MBAFCE9F0573.html>
Seq primer: T3.

FEATURES
SOURCE Location/Qualifiers

1..60
/organism="Brugia malayi"
/db_xref="taxon:6279"
/clone="AFCE9F05"
/clone_1lb="Brugia malayi adult female cDNA
(SAM96MLW-BMAF)"
/sex="female"
/dev_stage="adult"
/lab_host="Xrl-Blue MRF"
/note="Vector: Lambda uni-2AP XR; Site_1: Eco RI; Site_2:
Xho I; Lymphatic filarial nematode parasite of humans.
mRNA was prepared from approximately 50 adult females
isolated from the peritoneal cavity of jirds and
converted to double-stranded cDNA using reverse
transcriptase and oligo(dT) followed by RNase H and DNA
pol I. The library has 5 x 10⁶ independent recombinants
and the average insert size is ~900bp. The library was
constructed by Michelle Lizotte-Waniewski. The
library is available from Dr.S.A.Williams, email:
genomesmith.edu"

BASE COUNT 10 a 4 c 8 g 38 t
ORIGIN

Query Match 51.4%; Score 14.4; DB 29; Length 60;
Best Local Similarity 75.0%; Pred. No. 1.6e+04;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 atgccagcaaatcaacgtaaga 26
|| ||||| ||| |||
Db 25 ATACCAAGCAAAACAAATAGA 2

RESULT 13
AM057193 82 bp mRNA EST 29-SEP-1999
LOCUS ca02405.y1 C elegans fem3 Q23 S1 Caenorhabditis elegans cDNA 5'
DEFINITION similar to gb:R13H9.2 (ELEGANS); WP:R09C12.7 CE02603 ;, mRNA
sequence.
ACCESSION AM057193.1 GI:5932832
VERSION AM057193
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.

REFERENCE Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditiia; Rhabditiida;
Rhabditiina; Rhabditioidae; Rhabditiidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 82)
AUTHORS Ward,S., Smith,H., Clifton,S., Marra,M., Hillier,L., Kucaba,T.,
Bowers,Y., Person,B., Swaller,T., Steptoe,M., Gibbons,M.,
Harvey,N., Ritter,E., Jackson,Y., McCann,R., Waterston,R. and
Wilson,R.

TITLE Dofarizona-Washu C. elegans EST project
JOURNAL Unpublished (1999)
COMMENT On Dec 20, 1995 this sequence version replaced gi:1135389.

Contact: Samuel Ward, Ph.D.
Dofarizona-Washu C. elegans EST project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

FEATURES Email: est@wustl.edu
SOURCE Contact Harold Smith (hes@u.arizona.edu) for further information
relating to organism, libraries, or clone availability.
Seq primer: -40UP from Gibco.
Location/Qualifiers

1..82
/organism="Caenorhabditis elegans"

/db_xref="taxon:6239"
/clone_1lb="C. elegans fem3 Q23 S1"
/lab_host="DH5alpha cells"
/note="Vector: pBluescript II SK+; Site_1: XhoI; Site_2:
NotI; This C elegans library was made from fem-3(q23)
worms (produce only sperm at 25 C). cDNA was generated
via oligo (dT) priming and directionally cloned into
pBluescript II SK+ vector with a modified polylinker.
This library was subtracted with a second fem-1(hc17)
(produce only oocytes at 25 C) library to enrich this
library for sperm specific genes. First strand cDNA
synthesis was primed with a NotI-15T oligo (sequence
5'-CAGTACGTCGATCGCAGCGCCGCTTTTCTTTT-3'). After
2nd strand synthesis, XhoI-EcoRI adaptors (stratagene
#901120) were ligated, and cDNA was digested with NotI and
ligated into XhoI-NotI digested vector."

BASE COUNT 20 a 22 c 22 g 18 t
ORIGIN

Query Match 51.4%; Score 14.4; DB 64; Length 82;
Best Local Similarity 75.0%; Pred. No. 1.6e+04;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 agatgccagcaaatcaacgtaaa 24
|| ||||| ||| |||
Db 9 AGCTGCCAAGCAATTCGCCGCTTA 32

RESULT 14
AA916294 87 bp mRNA EST 17-JUN-1998
LOCUS on22b10.s1 NCI-CGAP.Lu5 Homo sapiens cDNA clone IMAGE:1557403 3'
DEFINITION similar to gb:X61123_rnal BTG1 PROTEIN (HUMAN);, mRNA sequence.
ACCESSION AA916294
VERSION AA916294.1 GI:3055686
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 87)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On May 5, 1995 this sequence version replaced gi:797728.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILM at:
www-bio.llnl.gov/dbfp/image/image.html

Trice considered overall poor quality
Insert Length: 990 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES Location/Qualifiers
SOURCE 1..87

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:1557403"
/clone_1lb="NCI-CGAP.Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from

neuroendocrine lung carcinoid, and was then primed with a
Not I - oligo(dT) primer. Double-stranded cDNA was ligated
to Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library is normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT

48 a 13 c 11 g 15 t

Db 44 ATACCAAGCAAAACAAAATTAGA 21

Search completed: June 4, 2000, 13:53:12
Job time: 20936 sec

Query Match 51.4%; Score 14.4; DB 40; Length 87;
Best Local Similarity 75.0%; Pred. No. 1.6e+04;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 agatgccagcaatcaacgctaa 24
1 | | | | | | | | | | | | | | | |
Db 3 AAATGCCAAGAAAGAAACGATTA 26

RESULT 15

AI784864 94 bp mRNA EST 02-JUL-1999
LOCUS SWAMC30B03SK Brugia malayi adult male cDNA (SAW94NL-BMAM) Brugia
DEFINITION malayi cDNA clone SWAMC30B03 5', mRNA sequence.

ACCESSION AI784864
VERSION AI784864.1 GI:5332489
KEYWORDS EST.
SOURCE Brugia malayi.
ORGANISM Brugia malayi.

Eukaryota; Metazoa; Nematoda; Secernentea; Spirurida;
Filarioidea; Onchocercidae; Brugia.

REFERENCE 1 (bases 1 to 94)
AUTHORS Williams, S.A.
TITLE Genes expressed in adult males of Brugia malayi
JOURNAL Unpublished (1995)
COMMENT On Jun 22, 1998 this sequence version replaced gi:3246848.

Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith
College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genomesmith.edu

FEATURES

Seq primer: pairescript SK.
Location/Qualifiers
1..94

/organism="Brugia malayi"
/strain="TRS Labs"
/db_xref="taxon:6279"
/clone="SWAMC30B03"
/lab_host="X11-Blue MRP"
/note="Vector: lambda Unizap XR; Site_1: EcoR I; Site_2:
Xho I; Lymphatic filarial nematode parasite of humans.
mRNA was prepared from adult males of Brugia malayi
isolated from jirds and converted to double stranded cDNA
using reverse transcriptase and oligo(dT) followed by
Rnase H and DNAPol I. The library had 4.6 x 10E6
independent recombinants and average insert size was 800
base pairs. The library was constructed by Noelle Ling.
The library is available from Dr. S.A. Williams, email
genomesmith.edu."

BASE COUNT 19 a 4 c 9 g 62 t

ORIGIN

Query Match 51.4%; Score 14.4; DB 60; Length 94;
Best Local Similarity 75.0%; Pred. No. 1.7e+04;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 atgcaagcaatcaacgctaa 26
1 | | | | | | | | | | | | | | | |

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:41 ; Search time 244.64 Seconds
(Without alignments)
14.877 Million cell updates/sec

Title: US-09-164-714-9

Perfect score: 28
Sequence: 1 agatgcgaagcaatcaacggtagaac 28

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 6492525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: /cgn2_Patents_NA:*
2: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5C.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/5D.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/6.COMB.seq:*
7: /cgn2_6/ptodata/2/ina/PCITUS.COMB.seq:*
7: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	57.1	79	2	US-08-446-908-8
2	16	57.1	79	2	US-08-231-205A-8
3	16	57.1	79	4	US-08-871-161-8
4	15.8	56.4	54	4	US-08-649-991-10
5	15.8	56.4	54	4	US-08-649-991-11
6	15.8	56.4	54	4	US-08-649-991-12
7	15.8	56.4	54	4	US-08-649-991-13
8	15.8	56.4	54	4	US-08-649-991-14
9	15.8	56.4	54	4	US-08-649-991-15
10	15.8	56.4	54	4	US-08-649-991-16
11	15.8	56.4	54	4	US-08-649-991-17
12	15.8	56.4	54	4	US-08-649-991-18
13	15.8	56.4	54	4	US-08-649-991-20
14	15.8	56.4	54	4	US-08-649-991-22
15	15.8	56.4	54	4	US-08-649-991-23
16	15.8	56.4	54	4	US-08-649-991-25
17	15.8	56.4	54	4	US-08-649-991-26
18	15.8	56.4	54	4	US-08-649-991-27
19	15.8	56.4	54	4	US-08-649-991-119
20	15.8	56.4	54	4	US-08-649-991-120
21	15.8	56.4	54	4	US-08-649-991-121
22	15.8	56.4	54	4	US-08-649-991-122
23	15.8	56.4	54	4	US-08-649-991-123
24	15.8	56.4	54	4	US-08-649-991-124
25	15.8	56.4	54	4	US-08-649-991-125
26	15.8	56.4	54	4	US-08-649-991-126
27	15.8	56.4	54	4	US-08-649-991-128

28	15.8	56.4	54	4	US-08-649-991-129	Sequence 129, App
29	15	53.6	31	1	US-08-390-850-43	Sequence 43, Appl
30	15	53.6	31	2	US-08-435-634-43	Sequence 43, Appl
31	14.8	52.9	53	3	US-08-811-492-120	Sequence 120, App
32	14.8	52.9	54	4	US-08-649-991-21	Sequence 21, Appl
33	14.8	52.9	54	4	US-08-649-991-24	Sequence 24, Appl
34	14.8	52.9	54	4	US-08-649-991-31	Sequence 31, Appl
35	14.8	52.9	54	6	PCR-US94-09700-35	Sequence 35, Appl
36	14.8	52.9	65	2	US-08-273-594-29	Sequence 29, Appl
37	14.6	52.1	69	3	US-07-916-098A-26	Sequence 26, Appl
38	14.6	52.1	69	3	US-07-916-098A-31	Sequence 31, Appl
39	14.2	50.7	50	4	US-08-850-049-5	Sequence 5, Appl
40	14.2	50.7	50	4	US-08-050-478-5	Sequence 5, Appl
41	14.2	50.7	54	4	US-08-649-991-9	Sequence 9, Appl
42	14.2	50.7	54	4	US-08-649-991-19	Sequence 19, Appl
43	14.2	50.7	57	7	5514566-17	Patent No. 5514566
44	14.2	50.7	99	2	US-08-463-115-69	Sequence 69, Appl
45	14.2	50.7	99	2	US-08-465-388-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1
US-08-446-908-8
Sequence 8, Application US/08446908
Patent No. 5705149
GENERAL INFORMATION:
APPLICANT: Namen, Anthony E.
APPLICANT: Goodwin, Raymond G.
APPLICANT: Lupton, Stephen D.
APPLICANT: Mochizuki, Diane Y.
TITLE OF INVENTION: Interleukin-7 and Antibodies Reactive
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.1
SOFTWARE: Microsoft Word, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,908
FILING DATE: 22-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/231,205
FILING DATE: 21-APR-1994
APPLICATION NUMBER: US 07/957,649
FILING DATE: 06-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,438
FILING DATE: 13-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/255,209
FILING DATE: 07-OCT-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/113,566
FILING DATE: 26-OCT-1987
ATTORNEY/AGENT INFORMATION:
NAME: Seese, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2104-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 8:

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 19..60
; US-08-446-908-8

Query Match          57.1%; Score 16; DB 2; Length 79;
Best Local Similarity 79.2%; Pred. No. 50;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      2  gatgccagcaatcacggtgaag 25
Db      34  GATGACAAGGAATCGAAGGTAGG 57

RESULT 2
; Sequence 8, Application US/08231205A
; Patent No. 571585
; GENERAL INFORMATION:
; APPLICANT: Namen, Anthony E.
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Lupton, Stephen D.
; APPLICANT: Mochizuki, Diane Y.
; TITLE OF INVENTION: Interleukin-7 and Antibodies Reactive
; TITLE OF INVENTION: Therewith
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Apple Macintosh
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,205A
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/957,649
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,438
; FILING DATE: 13-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/255,209
; FILING DATE: 07-OCT-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/113,566
; FILING DATE: 26-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2104-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```

;
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 19..60
; US-08-231-205A-8

Query Match          57.1%; Score 16; DB 2; Length 79;
Best Local Similarity 79.2%; Pred. No. 50;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      2  gatgccagcaatcacggtgaag 25
Db      34  GATGACAAGGAATCGAAGGTAGG 57

RESULT 3
; Sequence 8, Application US/08871161
; Patent No. 5965122
; GENERAL INFORMATION:
; APPLICANT: Namen, Anthony E.
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Lupton, Stephen D.
; APPLICANT: Mochizuki, Diane Y.
; TITLE OF INVENTION: Interleukin-7 and Antibodies Reactive
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Apple Macintosh
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,161
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,908
; FILING DATE: 22-MAY-1995
; APPLICATION NUMBER: US 08/231,205
; FILING DATE: 21-APR-1994
; APPLICATION NUMBER: US 07/957,649
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,438
; FILING DATE: 13-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/255,209
; FILING DATE: 07-OCT-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/113,566
; FILING DATE: 26-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2104-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```


CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ. ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-12

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacggtaga 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGCAAAACAAAGTAAGAA 47

RESULT 7
US-08-649-991-13
Sequence 13, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Narwa, Remy
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995

ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ. ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-13

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacggtaga 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGCAAAACAAAGTAAGAA 47

RESULT 8
US-08-649-991-14
Sequence 14, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Narwa, Remy
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ. ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-14

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgcacgaacaaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGACGCAAAACAAAGTAGAA 47

RESULT 9

US-08-649-991-15
; Sequence 15, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-649-991-15

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgcacgaacaaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGACGCAAAACAAAGTAGAA 47

RESULT 10
US-08-649-991-16
; Sequence 16, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy

APPLICANT: Roques, Pierre
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-649-991-16

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgcacgaacaaatcaacgtaagaa 27
||| | ||||| ||| |||||
Db 21 AGAGGAAGACGCAAAACAAAGTAGAA 47

RESULT 11
US-08-649-991-17
; Sequence 17, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy

; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-649-991-17

Query Match          56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaagcaaatcaacgtaagaa 27
    ||| | ||||| ||| ||||| |||
DB 21 AGAGGAAGACCAAAACAAAGTAAGAA 47

RESULT 12
US-08-649-991-18
; Sequence 18, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; APPLICANT: Roques, Pierre
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
```

```

; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-649-991-18

Query Match          56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaagcaaatcaacgtaagaa 27
    ||| | ||||| ||| ||||| |||
DB 21 AGAGGAAGACCAAAACAAAGTAAGAA 47

RESULT 13
US-08-649-991-20
; Sequence 20, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; APPLICANT: Roques, Pierre
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-649-991-20

Query Match          56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaagcaaatcaacgtaagaa 27
    ||| | ||||| ||| ||||| |||
```

Db 21 AGAGGAGAGCAAAACAAGTAGAA 47

RESULT 14

US-08-649-991-22

; Sequence 22, Application US/08649991

; Patent No. 5919462

; GENERAL INFORMATION:

; APPLICANT: Narwa, Remy

; APPLICANT: Roques, Pierre

; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE

; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR

; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF

; TITLE OF INVENTION: MATERNOCERATL TRANSMISSION OF HIV-1

; NUMBER OF SEQUENCES: 130

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP

; STREET: 1800 M Street, N.W.

; City: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20036-5869

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/649,991

; FILING DATE: 17-MAY-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: FR 9505914

; FILING DATE: 18-MAY-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Adler, Reid G.

; REGISTRATION NUMBER: 30,988

; REFERENCE/DOCKET NUMBER: ORES-5003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-467-7000

; TELEFAX: 202-467-7176

; INFORMATION FOR SEQ ID NO: 22:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 54 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; US-08-649-991-22

Query Match

Best Local Similarity 56.4%; Score 15.8; DB 4; Length 54;

Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacgtaagaa 27

Db 21 AGAGGAGAGCAAAACAAGTAGAA 47

RESULT 15

US-08-649-991-23

; Sequence 23, Application US/08649991

; Patent No. 5919462

; GENERAL INFORMATION:

; APPLICANT: Narwa, Remy

; APPLICANT: Roques, Pierre

; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE

; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR

; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF

; TITLE OF INVENTION: MATERNOCERATL TRANSMISSION OF HIV-1

; NUMBER OF SEQUENCES: 130

; CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
City: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20036-5869

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/649,991

; FILING DATE: 17-MAY-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: FR 9505914

; FILING DATE: 18-MAY-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Adler, Reid G.

; REGISTRATION NUMBER: 30,988

; REFERENCE/DOCKET NUMBER: ORES-5003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-467-7000

; TELEFAX: 202-467-7176

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 54 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; US-08-649-991-23

Query Match

Best Local Similarity 56.4%; Score 15.8; DB 4; Length 54;

Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agatgccaaagcaatcaacgtaagaa 27

Db 21 AGAGGAGAGCAAAACAAGTAGAA 47

Search completed: June 4, 2000, 16:09:42
Job time: 28061 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:33 ; Search time 1236.38 Seconds

(without alignments)
-22.031 Million cell updates/sec

Title: US-09-164-714-10

Perfect score: 28

Sequence: 1 gttctacgtgattgcttgcgtgcacatc 28

Scoring table: IDENTITY_NUC

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0

Post-processing: Minimum Match 0%

Database : Listing first 45 summaries

1: gb_bal:*
2: gb_bal:*
3: gb_bal:*
4: gb_bal:*
5: gb_bal:*
6: gb_bal:*
7: gb_bal:*
8: gb_bal:*
9: gb_bal:*
10: gb_bal:*
11: gb_bal:*
12: gb_bal:*
13: gb_bal:*
14: gb_bal:*
15: gb_bal:*
16: gb_bal:*
17: gb_bal:*
18: gb_bal:*
19: gb_bal:*
20: gb_bal:*
21: gb_bal:*
22: gb_bal:*
23: gb_bal:*
24: gb_bal:*
25: gb_bal:*
26: gb_bal:*
27: gb_bal:*
28: gb_bal:*
29: gb_bal:*
30: gb_bal:*
31: gb_bal:*
32: gb_bal:*
33: gb_bal:*
34: gb_bal:*
35: gb_bal:*
36: gb_bal:*
37: gb_bal:*
38: gb_bal:*
39: gb_bal:*
40: gb_bal:*
41: gb_bal:*
42: gb_bal:*
43: gb_bal:*
44: gb_bal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	57.1	79	107076	107076 Sequence 2
2	16	57.1	79	183771	183771 Sequence 8
3	15.8	56.4	54	A57944	A57944 Sequence 10
4	15.8	56.4	54	A57945	A57945 Sequence 11
5	15.8	56.4	54	A57946	A57946 Sequence 12
6	15.8	56.4	54	A57947	A57947 Sequence 13
7	15.8	56.4	54	A57948	A57948 Sequence 14
8	15.8	56.4	54	A57949	A57949 Sequence 15
9	15.8	56.4	54	A57950	A57950 Sequence 16
10	15.8	56.4	54	A57951	A57951 Sequence 17
11	15.8	56.4	54	A57952	A57952 Sequence 18
12	15.8	56.4	54	A57953	A57953 Sequence 19
13	15.8	56.4	54	A57954	A57954 Sequence 20
14	15.8	56.4	54	A57955	A57955 Sequence 21
15	15.8	56.4	54	A57956	A57956 Sequence 22
16	15.8	56.4	54	A57957	A57957 Sequence 23
17	15.8	56.4	54	A57958	A57958 Sequence 24
18	15.8	56.4	54	A57959	A57959 Sequence 25
19	15.8	56.4	54	A57960	A57960 Sequence 26
20	15.8	56.4	54	A57961	A57961 Sequence 27
21	15.8	56.4	54	E05823	E05823 DNA encod1n
22	15.8	56.4	54	I37030	I37030 Sequence 43
23	15.8	56.4	54	I37030	I37030 Sequence 43
24	15.8	56.4	54	I37030	I37030 Sequence 43
25	15.8	56.4	54	I37030	I37030 Sequence 43
26	15.8	56.4	54	I37030	I37030 Sequence 43
27	15.8	56.4	54	I37030	I37030 Sequence 43
28	15.8	56.4	54	I37030	I37030 Sequence 43
29	15.8	56.4	54	I37030	I37030 Sequence 43
30	15.8	56.4	54	I37030	I37030 Sequence 43
31	15.8	56.4	54	I37030	I37030 Sequence 43
32	15.8	56.4	54	I37030	I37030 Sequence 43
33	15.8	56.4	54	I37030	I37030 Sequence 43
34	15.8	56.4	54	I37030	I37030 Sequence 43
35	15.8	56.4	54	I37030	I37030 Sequence 43
36	15.8	56.4	54	I37030	I37030 Sequence 43
37	15.8	56.4	54	I37030	I37030 Sequence 43
38	15.8	56.4	54	I37030	I37030 Sequence 43
39	15.8	56.4	54	I37030	I37030 Sequence 43
40	15.8	56.4	54	I37030	I37030 Sequence 43
41	15.8	56.4	54	I37030	I37030 Sequence 43
42	15.8	56.4	54	I37030	I37030 Sequence 43
43	15.8	56.4	54	I37030	I37030 Sequence 43
44	15.8	56.4	54	I37030	I37030 Sequence 43
45	15.8	56.4	54	I37030	I37030 Sequence 43

ALIGNMENTS

RESULT 1
LOCUS 107076 79 bp PAT 02-DEC-1994
DEFINITION Sequence 2 from Patent EP 0314415.
ACCESSION 107076
VERSION 107076.1 GI:590362
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 79)
AUTHORS Namen,A.E., Goodwin,R.G., Lupton,S.D., Mochizuki,D.Y., Price,V.L.
TITLE Interleukin-7
JOURNAL Patent: EP 0314415-A2 2 03-MAY-1989;
FEATURES Location/Qualifiers
source 1..79
BASE COUNT 27 a 18 c 22 g 12 t
ORIGIN

Query Match 57.1%; Score 16; DB 5; Length 79;
Best Local Similarity 79.2%; Pred. No. 5.8e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 ctaccgtgattgcttgcgcac 27
Db 57 CCTACCTTGATTCTGCTGTCATC 34

RESULT 2
LOCUS 183771 79 bp DNA PAT 10-AUG-1998
DEFINITION Sequence 8 from patent US 5714585.
ACCESSION 183771
VERSION 183771.1 GI:3407301
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 79)
AUTHORS Namen,A.E., Goodwin,R.G., Lupton,S.D. and Mochizuki,D.Y.
TITLE Antibodies that are immunoreactive with interleukin-7
JOURNAL Patent: US 5714585-A 8 03-FEB-1998;
FEATURES Location/Qualifiers
source 1..79
BASE COUNT 27 a 18 c 22 g 12 t
ORIGIN

Query Match 57.1%; Score 16; DB 5; Length 79;
Best Local Similarity 79.2%; Pred. No. 5.8e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 ctaccgtgattgcttgcgcac 27
Db 57 CCTACCTTGATTCTGCTGTCATC 34

RESULT 3
LOCUS A57944 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 10 from Patent EP0743364.
ACCESSION A57944
VERSION A57944.1 GI:3713714
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)

AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding of fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 10 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 31 a 4 c 13 g 6 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttctaccgttattgcttgcacatc 28
Db 47 TTCCTACTTTGCTTTCCTCTCCTCT 21

RESULT 4
LOCUS A57945 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 11 from Patent EP0743364.
ACCESSION A57945
VERSION A57945.1 GI:3713715
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding of fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 11 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 4 c 15 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttctaccgttattgcttgcacatc 28
Db 47 TTCCTACTTTGCTTTCCTCTCCTCT 21

RESULT 5
LOCUS A57946 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 12 from Patent EP0743364.
ACCESSION A57946
VERSION A57946.1 GI:3713716
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding of fragments and their application as reactives for risk evaluation of

JOURNAL HIV-1 mother-foetal transmission
Patent: EP 0743364-A 12-20-NOV-1996;
COMMISSARIAT ENERGIE ATOMIQUE (FR)
Other publication FR 2734281 961122.
FEATURES
source 1..54
/db_xref="taxon:32644"
Location/Qualifiers

BASE COUNT 31 a 4 c 14 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttctacgctgattgctgcatct 28
||||| ||| ||||| |||
Db 47 TTCTACTTTGTGTTGCTCTCTCT 21

RESULT 6
A57947/c 54 bp DNA PAT 05-MAR-1998
LOCUS A57947 Sequence 13 from Patent EP0743364.
DEFINITION A57947
ACCESSION A57947
VERSION A57947.1 GI:3713717
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 13-20-NOV-1996;
COMMISSARIAT ENERGIE ATOMIQUE (FR)
COMMENT Other publication FR 2734281 961122.
FEATURES
source 1..54
/db_xref="taxon:32644"
Location/Qualifiers

BASE COUNT 31 a 4 c 13 g 6 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttctacgctgattgctgcatct 28
||||| ||| ||||| |||
Db 47 TTCTACTTTGTGTTGCTCTCTCT 21

RESULT 7
A57948/c 54 bp DNA PAT 05-MAR-1998
LOCUS A57948 Sequence 14 from Patent EP0743364.
DEFINITION A57948
ACCESSION A57948
VERSION A57948.1 GI:3713718
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 14-20-NOV-1996;
COMMISSARIAT ENERGIE ATOMIQUE (FR)

COMMENT Other publication FR 2734281 961122.
FEATURES
source 1..54
/db_xref="taxon:32644"
Location/Qualifiers

BASE COUNT 32 a 5 c 12 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttctacgctgattgctgcatct 28
||||| ||| ||||| |||
Db 47 TTCTACTTTGTGTTGCTCTCTCT 21

RESULT 8
A57949/c 54 bp DNA PAT 05-MAR-1998
LOCUS A57949 Sequence 15 from Patent EP0743364.
DEFINITION A57949
ACCESSION A57949
VERSION A57949.1 GI:3713719
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 15-20-NOV-1996;
COMMISSARIAT ENERGIE ATOMIQUE (FR)
COMMENT Other publication FR 2734281 961122.
FEATURES
source 1..54
/db_xref="taxon:32644"
Location/Qualifiers

BASE COUNT 30 a 5 c 14 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttctacgctgattgctgcatct 28
||||| ||| ||||| |||
Db 47 TTCTACTTTGTGTTGCTCTCTCT 21

RESULT 9
A57950/c 54 bp DNA PAT 05-MAR-1998
LOCUS A57950 Sequence 16 from Patent EP0743364.
DEFINITION A57950
ACCESSION A57950
VERSION A57950.1 GI:3713720
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 16-20-NOV-1996;
COMMISSARIAT ENERGIE ATOMIQUE (FR)
COMMENT Other publication FR 2734281 961122.
FEATURES
source 1..54
Location/Qualifiers

BASE COUNT 30 a 5 c 14 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ttctacgcttgattgctgcacatc 28
||||| ||| ||| ||| ||| |||
Db 47 TTCTTACTTTTGTGCTCTCTCT 21

RESULT 10
A57951/c 54 bp DNA PAT 05-MAR-1998
LOCUS A57951 Sequence 17 from Patent EP0743364.
DEFINITION A57951
ACCESSION A57951
VERSION A57951.1 GI:3713721
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 17 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 31 a 4 c 14 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ttctacgcttgattgctgcacatc 28
||||| ||| ||| ||| ||| |||
Db 47 TTCTTACTTTTGTGCTCTCTCT 21

RESULT 11
A57952/c 54 bp DNA PAT 05-MAR-1998
LOCUS A57952 Sequence 18 from Patent EP0743364.
DEFINITION A57952
ACCESSION A57952
VERSION A57952.1 GI:3713722
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 18 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 29 a 5 c 15 g 4 t 1 others

ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ttctacgcttgattgctgcacatc 28
||||| ||| ||| ||| ||| |||
Db 47 TTCTTACTTTTGTGCTCTCTCT 21

RESULT 12
A57954/c 54 bp DNA PAT 05-MAR-1998
LOCUS A57954 Sequence 20 from Patent EP0743364.
DEFINITION A57954
ACCESSION A57954
VERSION A57954.1 GI:3713724
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 20 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 30 a 4 c 15 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ttctacgcttgattgctgcacatc 28
||||| ||| ||| ||| ||| |||
Db 47 TTCTTACTTTTGTGCTCTCTCT 21

RESULT 13
A57956/c 54 bp DNA PAT 05-MAR-1998
LOCUS A57956 Sequence 22 from Patent EP0743364.
DEFINITION A57956
ACCESSION A57956
VERSION A57956.1 GI:3713726
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 22 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 30 a 6 c 14 g 4 t

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ttcttaccgttgattgcttgcattc 28
||||| ||| ||||| | |||
Db 47 TTCTTACTTTGTGTGGCTCTCTCT 21

QY 2 ttcttaccgttgattgcttgcattc 28
||||| ||| ||||| | |||
Db 47 TTCTTACTTTGTGTGGCTCTCTCT 21

Search completed: June 4, 2000, 16:05:33
Job time: 27885 sec

RESULT 14
A57957/c 54 bp DNA PAT 05-MAR-1998
LOCUS A57957 Sequence 23 from Patent EP0743364.
DEFINITION A57957
ACCESSION A57957
VERSION A57957.1 GI:3713727
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 23 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source 1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 6 c 13 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ttcttaccgttgattgcttgcattc 28
||||| ||| ||||| | |||
Db 47 TTCTTACTTTGTGTGGCTCTCTCT 21

RESULT 15
A57959/c 54 bp DNA PAT 05-MAR-1998
LOCUS A57959 Sequence 25 from Patent EP0743364.
DEFINITION A57959
ACCESSION A57959
VERSION A57959.1 GI:3713729
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 25 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source 1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 4 c 15 g 5 t
ORIGIN

Query Match 56.4%; Score 15.8; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 7.2e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:08 ; Search time 322.35 Seconds
(without alignments)
21.732 Million cell updates/sec

Title: US-09-164-714-10

Perfect score: 28
Sequence: 1 gtcttaccgttgattgcttgcatct 28

Scoring table: IDENTITY_NNC
Gapop 10.0 , Gapext 1.0

Searched: 311585 segs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	15.8	56.4	54	1 T43658	HIV-1 matrix prote
C 2	15.8	56.4	54	1 T43659	HIV-1 matrix prote
C 3	15.8	56.4	54	1 T43642	HIV-1 matrix prote
C 4	15.8	56.4	54	1 T43643	HIV-1 matrix prote
C 5	15.8	56.4	54	1 T43644	HIV-1 matrix prote
C 6	15.8	56.4	54	1 T43645	HIV-1 matrix prote
C 7	15.8	56.4	54	1 T43646	HIV-1 matrix prote
C 8	15.8	56.4	54	1 T43647	HIV-1 matrix prote
C 9	15.8	56.4	54	1 T43648	HIV-1 matrix prote
C 10	15.8	56.4	54	1 T43649	HIV-1 matrix prote
C 11	15.8	56.4	54	1 T43650	HIV-1 matrix prote
C 12	15.8	56.4	54	1 T43652	HIV-1 matrix prote
C 13	15.8	56.4	54	1 T43654	HIV-1 matrix prote
C 14	15.8	56.4	54	1 T43655	HIV-1 matrix prote
C 15	15.8	56.4	54	1 T43657	HIV-1 matrix prote
C 16	15.8	56.4	92	1 O53311	Hybrid antigen pro
C 17	15.2	54.3	38	1 N94509	Probe for N-term
C 18	15.2	54.3	38	1 V83523	PCR primer used to
C 19	15.2	54.3	38	1 V83525	PCR primer used to
C 20	15.2	54.3	97	1 T19416	Human gene signatu
C 21	15	53.6	31	1 O93504	Human stromelysin
C 22	14.8	52.9	47	1 X52594	Human genome biall
C 23	14.8	52.9	47	1 X52596	Human genome biall
C 24	14.8	52.9	47	1 X52546	Human genome biall
C 25	14.8	52.9	47	1 X52544	Human genome biall
C 26	14.8	52.9	53	1 V68220	PCR primer used to
C 27	14.8	52.9	54	1 O86781	SDI-1 mimetic prim
C 28	14.8	52.9	54	1 O74004	SDI-1 mimetic PCR
C 29	14.8	52.9	54	1 T43663	HIV-1 matrix prote
C 30	14.8	52.9	54	1 T43663	HIV-1 matrix prote
C 31	14.8	52.9	54	1 T43656	HIV-1 matrix prote
C 32	14.6	52.1	43	1 T01299	PCR primer Oligo D
C 33	14.6	52.1	69	1 O30897	Primer 312-69. New
C 34	14.6	52.1	69	1 O30892	Primer 312-64. New

35	14.6	52.1	100	1 T42735
36	14.4	51.4	60	1 T62581
C 37	14.2	50.7	47	1 V68272
C 38	14.2	50.7	54	1 T43651
C 39	14.2	50.7	54	1 T43641
C 40	14.2	50.7	99	1 V10203
C 41	14.2	50.7	99	1 V12016
C 42	14	50.0	99	1 T30883
C 43	13.8	49.3	23	1 O88336
C 44	13.8	49.3	30	1 O88341
C 45	13.8	49.3	37	1 V39939

ALIGNMENTS

RESULT 1
T43658/c
ID T43658; standard; DNA; 54 BP.
AC T43658;
DT 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment 2754.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R. Roques P;
DR WPT: 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 27; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, GTA, CTA, GTC or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 30 A; 4 C; 15 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttcttaccgttgattgcttgcatct 28
DB 47 tttcttaccgttgattgcttgcatct 21

RESULT 2
T43659/c
ID T43659; standard; DNA; 54 BP.
AC T43659;
DT 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment 2826.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.

PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;
DR WPI: 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 27; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, GTA, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 Bp; 30 A; 4 C; 15 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 2 ttctaccgttgattgctgcatct 28
Db 47 TTCTACTTGTGTTTGCTCTCCCT 21

RESULT 3
T43642/C
ID T43642 standard; DNA: 54 Bp.
AC T43642.
DT 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment ARI.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN Ep-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;
DR WPI: 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 23; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, GTA, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 Bp; 31 A; 4 C; 13 G; 6 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 2 ttctaccgttgattgctgcatct 28
Db 47 TTCTACTTGTGTTTGCTCTCCCT 21

RESULT 4
T43643/C
ID T43643 standard; DNA: 54 Bp.
AC T43643.
DT 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment BOI.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
FH Key Location/Qualifiers
FT mat_peptide 1..54
FT /*tag= a
FT /product= BOI_peptide
FN Ep-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;
DR WPI: 96-507733/51.
DR P-PSDB; W06610.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 23; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, GTA, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 Bp; 30 A; 4 C; 15 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 2 ttctaccgttgattgctgcatct 28
Db 47 TTCTACTTGTGTTTGCTCTCCCT 21

RESULT 5
T43644/C
ID T43644 standard; DNA: 54 Bp.
AC T43644.
DT 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment DUM.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN Ep-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;
DR WPI: 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal

PT transmission of HIV-1 infection
PS Claim 3: Page 24; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 31 A; 4 C; 14 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttctaccgtgattgctgcgcattc 28
||||| ||| ||||| |||
DB 47 TTCTTACTTTTGTGTTGCTCTCTCTCT 21

RESULT 6
T43645/c
ID T43645 standard; DNA: 54 BP.
AC T43645;
DE 19-AUG-1997 (first entry)
DT HIV-1 matrix protein p17 gene fragment PAL.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.
PD 20-NOV-1996.
PR 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R. Roques P;
PI WPI: 96-507733/51.
DR Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3: Page 24; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 31 A; 4 C; 13 G; 6 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttctaccgtgattgctgcgcattc 28
||||| ||| ||||| |||
DB 47 TTCTTACTTTTGTGTTGCTCTCTCTCT 21

RESULT 7

T43646/c
ID T43646 standard; DNA: 54 BP.
AC T43646;
DE 19-AUG-1997 (first entry)
DT HIV-1 matrix protein p17 gene fragment RYO.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.
PD 20-NOV-1996.
PR 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R. Roques P;
PI WPI: 96-507733/51.
DR Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3: Page 24; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 32 A; 5 C; 12 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttctaccgtgattgctgcgcattc 28
||||| ||| ||||| |||
DB 47 TTCTTACTTTTGTGTTGCTCTCTCTCT 21

RESULT 8
T43647/c
ID T43647 standard; DNA: 54 BP.
AC T43647;
DE 19-AUG-1997 (first entry)
DT HIV-1 matrix protein p17 gene fragment FLO.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.
PD 20-NOV-1996.
PR 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R. Roques P;
PI WPI: 96-507733/51.
DR Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3: Page 24; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.

CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SO Sequence 54 BP; 30 A; 5 C; 14 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 2 ttctaccgttgattgctgcatcct 28
||||| ||| ||||| |||
Db 47 TTCTTACTTTGTGTTGGCTCTCCTCT 21

RESULT 9
T43648/C
ID T43648 standard; DNA: 54 BP.
AC T43648;
DE 19-AUG-1997 (first entry)
KW HIV-1 matrix protein p17 gene fragment 4501.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
FH Key Location/Qualifiers
FT mat_peptide 1..54
FT /tag= a
FT /product= 4501_peptide

EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;
DR WPI: 96-507733/51.
P-PSDB; W06611.

PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 24; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SO Sequence 54 BP; 30 A; 5 C; 14 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 2 ttctaccgttgattgctgcatcct 28
||||| ||| ||||| |||
Db 47 TTCTTACTTTGTGTTGGCTCTCCTCT 21

RESULT 10
T43649/C
ID T43649 standard; DNA: 54 BP.
AC T43649;
DE 19-AUG-1997 (first entry)
KW HIV-1 matrix protein p17 gene fragment pAL.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;

KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.

PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;
DR WPI: 96-507733/51.

PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 25; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SO Sequence 54 BP; 31 A; 4 C; 14 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 2 ttctaccgttgattgctgcatcct 28
||||| ||| ||||| |||
Db 47 TTCTTACTTTGTGTTGGCTCTCCTCT 21

RESULT 11
T43650/C
ID T43650 standard; DNA: 54 BP.
AC T43650;
DE 19-AUG-1997 (first entry)

DE HIV-1 matrix protein p17 gene fragment HAR.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.

OS Human immunodeficiency virus type 1.
FH Key Location/Qualifiers
FT mat_peptide 1..54
FT /tag= a
FT /product= HAR_peptide

EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;
DR WPI: 96-507733/51.
P-PSDB; W06612.

PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 25; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining

CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 29 A; 5 C; 15 G; 4 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ttctacgttgattgctgcacatc 28
DB 47 TTCTTACTTTGTTGCTTCCTCCTCCT 21

RESULT 12
T43652/C
ID T43652 standard; DNA; 54 BP.
AC T43652;

DE 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment CHET.
KM Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN EP-743364-A2.

PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

PI Narwa R. Roques P;
DR MPI: 96-507733/51.

PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 26; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GAG, GAC, GAT or AGA; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 30 A; 4 C; 15 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ttctacgttgattgctgcacatc 28
DB 47 TTCTTACTTTGTTGCTTCCTCCTCCT 21

RESULT 13
T43654/C
ID T43654 standard; DNA; 54 BP.
AC T43654;

DE 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment SIW.
KM Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN Location/Qualifiers
FH Key 1.54
FT mat_peptide
FT /tag= a
FT /product= SIW_peptide

PN EP-743364-A2.
PD 20-NOV-1996.
PF 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

PI Narwa R. Roques P;
DR MPI: 96-507733/51.
DR P-PSDB; W06614.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 26; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GAG, GAC, GAT or AGA; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 30 A; 6 C; 14 G; 4 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ttctacgttgattgctgcacatc 28
DB 47 TTCTTACTTTGTTGCTTCCTCCTCCT 21

RESULT 14
T43655/C
ID T43655 standard; DNA; 54 BP.
AC T43655;

DE 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment MOE.
KM Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN Location/Qualifiers
FH Key 1.54
FT mat_peptide
FT /tag= a
FT /product= MOE_peptide

PD EP-743364-A2.
PF 20-NOV-1996.
PR 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

PI Narwa R. Roques P;
DR MPI: 96-507733/51.
DR P-PSDB; W06615.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 26; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTA, GAG, GAC, GAT or AGA; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using

CC standard hybridisation or immuno assays. The presence of such sequences
 CC in maternal blood is strongly correlated with transmission of infection.
 SQ Sequence 54 Bp; 30 A; 6 C; 13 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
 Best Local Similarity 74.1%; Pred. No. 2.1e+02;
 Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 2 ttctaccgttgattgctgcacatc 28
 ||||| ||||| ||||| |||||
 Db 47 TTCTTACTTTTGTGCTCTCTCTCT 21

RESULT 15

T43657/c
 ID T43657 standard; DNA; 54 Bp.
 AC T43657;
 DT 19-AUG-1997 (first entry)
 DE HIV-1 matrix protein p17 gene fragment 4541.
 KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
 KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
 OS Human immunodeficiency virus type 1.
 PN EP-743364-A2.
 PD 20-NOV-1996.
 PR 17-MAY-1996; 401084.
 PR 18-MAY-1995; FR-005914.
 PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 PI Narwa R, Rogues P;
 DR WPI: 96-507733/51.
 PT Human immunodeficiency virus p17 gene fragments, derived proteins
 PT and antibodies - useful for assessing the risk of maternal
 PT transmission of HIV-1 infection
 PS Claim 3: Page 27; 46pp; French.
 CC This sequence is a specifically claimed example of 21-90 nucleotide
 CC long nucleic acid fragments, derived from the gene encoding part of the
 CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
 CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
 CC Y6 = ATA, TTA, CTA, GTA, GTC or ATG; Y7 = GAG or GAA; Y8 = GAA or
 CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
 CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
 CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
 CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
 CC The new nucleic acid fragments can be used as reagents for determining
 CC and assessing the risk of maternal-foetal transmission of HIV-1, using
 CC standard hybridisation or immuno assays. The presence of such sequences
 CC in maternal blood is strongly correlated with transmission of infection.
 SQ Sequence 54 Bp; 30 A; 4 C; 15 G; 5 T;

Query Match 56.4%; Score 15.8; DB 1; Length 54;
 Best Local Similarity 74.1%; Pred. No. 2.1e+02;
 Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 2 ttctaccgttgattgctgcacatc 28
 ||||| ||||| ||||| |||||
 Db 47 TTCTTACTTTTGTGCTCTCTCTCT 21

Search completed: June 4, 2000, 16:24:08
 Job time: 28904 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:12 ; Search time 4521.53 Seconds
(without alignments)
25.100 Million cell updates/sec

Title: US-09-164-714-10

Perfect score: 28
Sequence: 1 gtctacgctgattgtctgcatct 28

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

ESR:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_est48:*
83: gb_est49:*
84: gb_est50:*
85: gb_est51:*
86: em_est34:*
87: em_est35:*
88: em_est36:*
89: em_est37:*
90: gb_est52:*
91: gb_est53:*
92: gb_est54:*
93: gb_est55:*
94: gb_est56:*
95: em_est38:*
96: em_est39:*
97: em_est40:*
98: em_est41:*
99: em_est42:*
100: em_est43:*
101: em_est44:*
102: gb_est57:*
103: gb_est58:*
104: em_est39:*
105: gb_est59:*
106: gb_est60:*
107: gb_est61:*
108: gb_est62:*
109: gb_est63:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
C 1	15.2	54.3	62	38	AA746404	AA746404 nm62g01.s
C 2	15.2	54.3	63	38	AA746400	AA746400 nm62c07.s
C 3	15.2	54.3	72	30	AA244996	AA244996 nm29g07.r
C 4	15.2	54.3	75	38	AA746399	AA746399 nm62c05.s
C 5	15.2	54.3	75	38	AA746402	AA746402 nm62c11.s
C 6	15.2	54.3	83	37	AA681983	AA681983 vu76c07.r
C 7	15.2	54.3	89	20	D11978	D11978 HUM0512E03
C 8	15.2	54.3	94	20	D12000	D12000 HUM0513A08
C 9	15.2	53.6	84	83	AF088162	AF088162 Homo sapi
C 10	14.8	52.9	81	40	AA930201	AA930201 vs71e03.r
C 11	14.6	52.1	90	47	AI536907	AI536907 col4a09.x
C 12	14.4	51.4	60	29	AA161708	AA161708 MBAC95FO
C 13	14.4	51.4	82	64	AA057193	AA057193 ca02a05.y
C 14	14.4	51.4	87	40	AA916294	AA916294 on22b10.s
C 15	14.4	51.4	94	60	AI784864	AI784864 SRAMC30
C 16	14.4	51.4	99	42	AI082999	AI082999 SRAMC318
C 17	14.2	50.7	43	37	AA682184	AA682184 ah51h12.s
C 18	14.2	50.7	85	40	AA910924	AA910924 ok85e02.s
C 19	14.2	50.7	92	24	H99072	H99072 yv90h05.s1
C 20	14.2	50.0	52	48	AI584790	AI584790 fb83d08.y
C 21	14.2	50.0	73	33	AA398108	AA398108 zt58d08.s
C 22	14.2	50.0	82	62	AI906839	AI906839 RC-B126-
C 23	13.8	49.3	29	20	TI17517	TI17517 gsr m46 The
C 24	13.8	49.3	80	30	AA249355	AA249355 j1284.seq
C 25	13.8	49.3	85	30	AA231275	AA231275 mw32c03.r
C 26	13.8	49.3	100	23	HI9067	HI9067 ym51b08.r1
C 27	13.6	48.6	64	81	AA409557	AA409557 SMOVC3CAM
C 28	13.6	48.6	67	44	AI286661	AI286661 ub95902.r
C 29	13.6	48.6	76	35	C58767	C58767 C58767 Yuj1
C 30	13.6	48.6	76	62	AI915855	AI915855 w95a07.x
C 31	13.6	48.6	78	30	AA230217	AA230217 nc13f02.r
C 32	13.6	48.6	88	63	AI960312	AI960312 sc81g04.y
C 33	13.6	48.6	91	79	AM286463	AM286463 lgi 333.B
C 34	13.6	48.6	92	28	C20817	C20817 HDMS000486
C 35	13.6	48.6	94	31	AA291480	AA291480 zt40b04.s
C 36	13.6	48.6	96	45	AI365056	AI365056 qtl2b12.x
C 37	13.6	48.6	97	103	AQ248664	AQ248664 Fl3111-SP
C 38	13.6	48.6	99	24	H74881	H74881 226 Deletio
C 39	13.6	48.6	100	34	AA509111	AA509111 MBACX7D0
C 40	13.4	47.9	46	94	AA073659	AA073659 EP(2)2525
C 41	13.4	47.9	55	69	AM102242	AM102242 sds5b11.y
C 42	13.4	47.9	58	41	AU008799	AU008799 AU008799
C 43	13.4	47.9	59	63	AI988246	AI988246 sc98a02.y
C 44	13.4	47.9	64	37	AA715443	AA715443 nv53b08.r
C 45	13.4	47.9	67	39	AA905978	AA905978 oj89c03.s

ALIGNMENTS

RESULT 1
LOCUS AA746404/c 62 bp mRNA 27-JAN-1998
DEFINITION nm62g01.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1251169 3',
mRNA sequence.
ACCESSION AA746404
VERSION AA746404.1 GI:2786390
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 62)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,

Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdip/image/image.html
Insert Length: 471 Std Error: 0.00
Seq primer: -40m13 fwd. ER from Amersham
High quality sequence stop: 57.
Location/Qualifiers
1. 62
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1251169"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCGAAGGAGGAGGCGCCCTCAATTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pRT3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 25 a 11 c 17 g 9 t

Query Match 54.3%; Score 15.2; DB 38; Length 62;
Best Local Similarity 71.4%; Pred. No. 7.3e+03;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 gttctaccgttattcttgcattc 28
Db 42 GTATTAGCTTGAAGTCTCTGCATCT 15

RESULT 2
LOCUS AA746400/c 63 bp mRNA 27-JAN-1998
DEFINITION nm62c07.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1251180 3',
mRNA sequence.
ACCESSION AA746400
VERSION AA746400.1 GI:2786386
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 63)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CCAP clone distribution information can be

Matches	20:	Conservative	0:	Mismatches	8:	Indels	0:	Gaps	0:
QY	1	gtcttaccgttgattgctgcacatc	28						
Db	49	gtcttaccttttaattactcttctcatct	22						
RESULT	7								
LOCUS	D11978	89 bp	mRNA	EST	02-DEC-1992				
DEFINITION	HMM0S12E03	Liver HepG2 cell line.	Homo sapiens	cdNA clone	s12e03,				
ACCESSION	D11978								
VERSION	D11978.1	GI:2148809							
KEYWORDS	EST.								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;								
AUTHORS	Eutheria; Primates; Catarrhini; Homiidae; Homo.								
TITLE	1 (bases 1 to 89)								
JOURNAL	Okubo,K., Hori,N., Matoba,R., Niiyama,T., Fukushima,A., Kojima,Y.								
MEDLINE	and Matsubara,K.								
COMMENT	large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression								
FEATURES	Nature Genet. 2, 173-179 (1992)								
source	94258199								
CONTACT:	Kousaku Okubo, Naohiro Hori, Ryo Matoba, Toshiyuki Niiyama, Atsushi Fukushima, Yuko Kojima & Kenichi Matsubara								
INSTITUTE:	Institute for Molecular and Cellular Biology								
UNIVERSITY:	Osaka University								
LOCATION:	1-3 Yamada-Oka,Suita, Osaka 565, Japan.								
QUALIFIERS	1. .89								
ORGANISM="Homo sapiens"									
/db_xref="GDB:D0S8180E"									
/db_xref="taxon:9606"									
/clone="s12e03"									
/clone_lib="Liver HepG2 cell line."									
/lab_host="E.coli"									
/note="3'-directed regional cDNA library. Cleaved by MboI and transformed into E.coli."									
BASE COUNT	14 a	21 c	22 g	32 t					
ORIGIN									
Query Match		54.3%	Score 15.2:	DB 20:	Length 89:				
Best Local Similarity		71.4%:	Pred. NO.7.5e+03:						
Matches	20:	Conservative	0:	Mismatches	8:	Indels	0:	Gaps	0:
QY	1	gtcttaccgttgattgctgcacatc	28						
Db	46	gtcttaccgttgattgctgcacatc	73						
RESULT	8								
LOCUS	D12000	94 bp	mRNA	EST	02-DEC-1992				
DEFINITION	HMM0S13A08	Liver HepG2 cell line.	Homo sapiens	cdNA clone	s13a08,				
ACCESSION	D12000								
VERSION	D12000.1	GI:2148823							
KEYWORDS	EST.								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;								
AUTHORS	Eutheria; Primates; Catarrhini; Homiidae; Homo.								
TITLE	1 (bases 1 to 94)								
JOURNAL	Okubo,K., Hori,N., Matoba,R., Niiyama,T., Fukushima,A., Kojima,Y.								
MEDLINE	and Matsubara,K.								
COMMENT	large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression								
Nature Genet.	2, 173-179 (1992)								

```

COMMENT      Contact: Kousaku Okubo, Naohito Hori, Ryo Matoba, Toshiyuki
              Niiyama, Atsushi Fukushima, Yoko Kojima & Kenichi Matsubara
              Institute for Molecular and Cellular Biology
              Osaka University
              1-3 Yamada-oka, Suita, Osaka 565, Japan.

FEATURES
  source
      location/Qualifiers
          1..94
            /organism="Homo sapiens"
            /db_xref="GDB:1058194E"
            /db_xref="taxon:9606"
            /clone="s13a08"
            /clone_lib="Liver HepG2 cell line."
            /lab_host="E.coli"
            /note="3"-directed regional cDNA library. Cleaved by MboI
            and transformed into E.coli."
            1 others

BASE COUNT   18 a      21 c      21 g      33 t      1 others

ORIGIN
Query Match      54.3%; Score 15.2; DB 20; Length 94;
Best Local Similarity 71.4%; Pred. No. 7.5e+03;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      1 gttctaccgttgattctgtcgcattc 28
         ||||| | ||||| | ||||| |
Db      46 gtcctccctggttgattgctgcattt 73

RESULT  9
AF088162      84 bp      DNA      GSS      10-DEC-1998
LOCUS      Homo sapiens chromosome 11 clone 49D4Ec3 map 11q13, genomic survey
DEFINITION      sequence.
ACCESSION      AF088162
VERSION      AF088162.1 GI:3982872
KEYWORDS      GSS.
SOURCE      human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE      1 (bases 1 to 84)
AUTHORS      Gerhardt,D.S. and Zhu,S.
TITLE      A transcript map of an 800 kb region on human chromosome 11q13,
              part of the candidate region for SCA5 and BBS1
              Hum. Genet. (1999) In press
REFERENCE      2 (bases 1 to 84)
AUTHORS      Gerhardt,D.S. and Zhu,S.
TITLE      Direct Submission
JOURNAL      Submitted (31-AUG-1998) Department of Genetics, Washington
              University School of Medicine, 4566 Scott Avenue, Box 8232, St.
              Louis, MO 63110, USA

FEATURES
  source
      location/Qualifiers
          1..84
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /map="11q13"
            /clone="49D4Ec3"
            /chromosome="11"

BASE COUNT   24 a      17 c      10 g      33 t

ORIGIN
Query Match      53.6%; Score 15; DB 83; Length 84;
Best Local Similarity 78.3%; Pred. No. 9.1e+03;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2 ttctaccgttgattctgtcgc 24
         ||| | | |||| | |||||
Db      12 TTCGTCCTTTGATGCTGCTTGGC 34

RESULT  10
AA930201/C

```

LOCUS AA930201 81 bp mRNA EST 23-APR-1998
DEFINITION vs71e03.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:1151740 5' similar to gb:U15647_cds1 Mus musculus (MOUSE),,
mRNA sequence.
ACCESSION AA930201
VERSION AA930201.1 GI:3079794
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS 1 (bases 1 to 81)
Marras, M., Hille, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2286367.
Contact: Maria M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:624948
Seq primer: -28m13 rev1 ET from Amersham.
FEATURES
source Location/Qualifiers
1..81
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:1151740"
/clone_1lb="Stratagene mouse skin (#937313)"
/sex="females"
/tissue_type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: skin; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR Vector: ~5'
adaptor sequence: 5' GAATTCGGCAGCG 3' -3' adaptor
sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3'."
BASE COUNT 35 a 21 c 9 g 16 t
ORIGIN
Query Match 52.9%; Score 14.8; DB 40; Length 81;
Best Local Similarity 73.1%; Pred. No. 1.1e+04;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 3 tcttcacgttgattgctgcacatc 28
||||| ||||| ||||| ||||| |||||
Db 66 tcttcacggcttgattgcatctt 41
RESULT 11
LOCUS A1536907 90 bp mRNA EST 12-MAY-1999
DEFINITION t014a09.x1 NCI-CCAP_Ut2 Homo sapiens cDNA clone IMAGE:2179000 3',
mRNA sequence.
ACCESSION A1536907
VERSION A1536907.1 GI:4451042
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 90)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CCAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL On May 18, 1998 this sequence version replaced gi:3138691.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/btrp/image/image.html
Insert Length: 1143 Std Error: 0.00
Seq primer: -40up from Gibco
High quality sequence stop: 85
POLYA-No.
FEATURES
source Location/Qualifiers
1..90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2179000"
/clone_1lb="NCI-CCAP Ut2"
/tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: PCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.85 kb. Life Technologies catalog #:
11539-012"
BASE COUNT 34 a 16 c 13 g 27 t
ORIGIN
Query Match 52.1%; Score 14.6; DB 47; Length 90;
Best Local Similarity 81.0%; Pred. No. 1.4e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 2 tcttcacgttgattgctgc 22
||||| ||||| ||||| ||||| |||||
Db 37 tctcttaccgttgctgccttg 17
RESULT 12
LOCUS A161708 60 bp mRNA EST 17-DEC-1996
DEFINITION MBARCE9P05r3 Brugia malayi adult female cDNA (SAM06MLM-BNAF) Brugia
malayi cDNA clone AFCE9P05 5', mRNA sequence.
ACCESSION A161708
VERSION A161708.1 GI:1737540
KEYWORDS EST.
SOURCE Brugia malayi.
ORGANISM Brugia malayi
Eukaryota; Metazoa; Nematoda; Secernentea; Spirurida;
Filarioidea; Onchocercidae; Brugia.
REFERENCE
AUTHORS 1 (bases 1 to 60)
Blaxter, M.L., Waterfall, M., Daub, J., Lizotte, M., Baron, L. and
Jones, S.J.
Genes expressed in adult female Brugia malayi
Unpublished (1996)
JOURNAL On May 18, 1995 this sequence version replaced gi:811474.
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450

Email: mark.blaxter@ed.ac.uk
The ABI trace of this sequence can be viewed at
<http://www.sanger.ac.uk/brugia/ABC/MAFC9F05T3.html>
Seq primer: 73.

FEATURES
source Location/Qualifiers

1..60
/organism="Brugia malayi"
/db_xref="taxon:6279"
/clone="AFCE9F05"
/clone_1lb="Brugia malayi adult female cDNA
(SAM96MLM-BMAF)"
/sex="female"
/dev_stage="adult"
/lab_host="XLI-Blue MRF"
/note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
Xho I; Lymphatic filarial nematode parasite of humans.
mRNA was prepared from approximately 50 adult females
isolated from the peritoneal cavity of jirds and
converted to double-stranded cDNA using reverse
transcriptase and oligo(dT) followed by RNase H and DNA
pol I. The library has 5 x 10E6 independent recombinants
and the average insert size is ~900bp. The library was
constructed by Michelle Lizotte-Waniewski. The
library is available from Dr.S.A.Williams, email:
genome@emh.edu."

BASE COUNT 10 a 4 c 8 g 38 t

ORIGIN

Query Match 51.4%; Score 14.4; DB 29; Length 60;
Best Local Similarity 75.0%; Pred. No. 1.6e+04;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 tctaccgtgattgctgcatc 26
||| | ||| ||||| |||
Db 2 TCTAATTTTGTTCCTGCTGAT 25

RESULT 13
AM057193/C 82 bp mRNA EST 29-SEP-1999
LOCUS ca02a05.y1 C elegans fem3 Q23 S1 Caenorhabditis elegans cDNA 5'
DEFINITION similar to gb:R13H9.2 (ELEGENS); WP:F09C12.7 CE02603 ;, mRNA
sequence.

ACCESSION AM057193 GI:5932832
VERSION AM057193
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.

REFERENCE 1 (bases 1 to 82)
AUTHORS Ward,S., Smith,H., Clifton,S., Marra,M., Hillier,L., Kucaba,T.,
Rhabdilinea; Rhabdilinea; Rhabdilinea; Peloderinae; Caenorhabditis.
Pape,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Steptoe,M., Gibbons,M.,
Harvey,N., Ritzer,E., Jackson,Y., McCann,R., Waterston,R. and
Wilson,R.

TITLE UofArizona-Washu C. elegans EST project
JOURNAL Unpublished (1999)
COMMENT On Dec 20, 1995 this sequence version replaced gi:1135389.
CONTACT: Samuel Ward, Ph.D.
UofArizona-Washu C. elegans EST project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Contact Harold Smith (hes@u.arizona.edu) for further information
relating to organism, libraries, or clone availability.
Seq primer: -40UP from Gibco.

FEATURES
source 1..82
/organism="Caenorhabditis elegans"

/db_xref="taxon:6239"
/clone_1lb="C. elegans fem3 Q23 S1"
/lab_host="DH5alpha cells"
/note="Vector: pBluescript II SK+; Site_1: XhoI; Site_2:
NotI; This C elegans library was made from fem-3(q23)
worms (produce only sperm at 25 C). cDNA was generated
via oligo (dT) priming and directionally cloned into
pBluescript II SK+ vector with a modified polylinker.
This library was subtracted with a second fem-1(nc17)
(produce only oocytes at 25 C) library to enrich this
library for sperm specific genes. First strand cDNA
synthesis was primed with a NotI-15T oligo (sequence
5'-GACGATGTCATGATGCGCGCGCCGCTTTTCTTTT-3'). After
2nd strand synthesis, XhoI-EcoRI adaptors (Stratagene
#901120) were ligated, and cDNA was digested with NotI and
ligated into XhoI-NotI digested vector."

BASE COUNT 20 a 22 c 22 g 18 t

ORIGIN

Query Match 51.4%; Score 14.4; DB 64; Length 82;
Best Local Similarity 75.0%; Pred. No. 1.6e+04;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 ttaccgtgattgctgcatc 28
||| | ||| ||||| |||
Db 32 TTAACGGCGAATTCCTGCGAGCT 9

RESULT 14
AA916294/C 87 bp mRNA EST 17-JUN-1998
LOCUS on22b10.s1 NCI-CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1557403 3'
DEFINITION similar to gb:X61123_rnal BRG1 PROTEIN (HUMAN);, mRNA sequence.
ACCESSION AA916294
VERSION AA916294.1 GI:3055686
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 87)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On May 5, 1995 this sequence version replaced gi:797728.
CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1350
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA library preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINT at:
www.bio.llnl.gov/dbp/image/image.html

TTrace considered overall poor quality
Insert Length: 990 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES
source 1..87
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1557403"
/clone_1lb="NCI-CGAP_Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from

neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT713 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

Query Match	51.48;	Score 14.4;	DB 40;	Length 87;
Best Local Similarity	75.08;	Pred. No. 1.6e+04;		
Matches 18; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0;

```
QY      5 ttaccggtgatctgcttgcatct 28
          ||| ||| ||| ||| ||| |
Db     26 TAATCGTTCCTTCTTGCAATT 3
```

RESULT	15
LOCUS	A1784864
DEFINITION	A1784864 94 bp mRNA EST 02-JUL-1999 SMMACG30B03K Brugia malayi adult male cDNA (SAM94NL-BmaM) Brugia
ACCESSION	malayi_cDNA clone SMMACG30B03 5', mRNA sequence.
VERSION	A1784864
KEYWORDS	A1784864.1 GI:5332489
SOURCE	EST
ORGANISM	Brugia malayi. Brugia malayi

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (pages 1 to 94)	Williams, S. A.	Genes expressed in adult males of <i>Brugia malayi</i>	Unpublished (1995)	On Jun 22, 1998 this sequence version replaced gi:3246848.

source	1. .94	location/qualifiers	Seq primer: pbluescript SK.	Email: genome@smith.edu
FEATURES				

source

1. 94
/organism="Brugia malayi"
/strain="TRS Labs"
/db_xref="taxon:6279"
/clone="SMAMCAC30B03"
/lab_host="XLI-Blue MRF"
/note="Vector: lambda unizap XR; Site:1: EcoR I; Site:2:
Xho I; lymphatic filarial nematode parasite of humans.
mRNA was prepared from adult males of Brugia malayi
isolated from jirds and converted to double stranded cDNA
using reverse transcriptase and oligo(dT) followed by
Rnase H and Dnapol I. The library had 4.6 x 10E6
Independent recombinants and average insert size was 800
base pairs. The library was constructed by Noelle Ling.
The library is available from Dr. S.A. Williams, email:
genome@smith.edu."

19 a 4 c 9 g 62 t

BASE COUNT
ORIGIN

Query Match	51.48;	Score 14.4;	DB 60;	Length 94;
Best Local Similarity	75.08;	Pred. No. 1.7e+04;		
Matches 18; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0;

Qy 3 tctaccgtgattgcttgcacat 26
||| | ||| ||||| ||| ||

```
Db      21 TCIAATTTTGTTCCTGGTAT 44

Search completed: June  4, 2000, 13:53:12
Job time: 20936 sec
```

Search completed: June 4, 2000, 13:53:12
Job time: 20936 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:42 ; Search time 244.64 Seconds
(without alignments)
14.877 Million cell updates/sec

Title: US-09-164-714-10

Perfect score: 28

Sequence: 1 gtcttaccgttgattgtgtgcacatc 28

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/6.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/PCrUS.COMB.seq:*
7: /cgn2_6/ptodata/2/lna/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	57.1	79	2	US-08-446-908-8
2	16	57.1	79	2	US-08-231-205A-8
3	16	57.1	79	4	US-08-871-161-8
4	15.8	56.4	54	4	US-08-649-991-10
5	15.8	56.4	54	4	US-08-649-991-11
6	15.8	56.4	54	4	US-08-649-991-12
7	15.8	56.4	54	4	US-08-649-991-13
8	15.8	56.4	54	4	US-08-649-991-14
9	15.8	56.4	54	4	US-08-649-991-15
10	15.8	56.4	54	4	US-08-649-991-16
11	15.8	56.4	54	4	US-08-649-991-17
12	15.8	56.4	54	4	US-08-649-991-18
13	15.8	56.4	54	4	US-08-649-991-19
14	15.8	56.4	54	4	US-08-649-991-20
15	15.8	56.4	54	4	US-08-649-991-21
16	15.8	56.4	54	4	US-08-649-991-22
17	15.8	56.4	54	4	US-08-649-991-23
18	15.8	56.4	54	4	US-08-649-991-24
19	15.8	56.4	54	4	US-08-649-991-25
20	15.8	56.4	54	4	US-08-649-991-26
21	15.8	56.4	54	4	US-08-649-991-27
22	15.8	56.4	54	4	US-08-649-991-28
23	15.8	56.4	54	4	US-08-649-991-29
24	15.8	56.4	54	4	US-08-649-991-30
25	15.8	56.4	54	4	US-08-649-991-31
26	15.8	56.4	54	4	US-08-649-991-32
27	15.8	56.4	54	4	US-08-649-991-33

c	28	15.8	56.4	54	4	US-08-649-991-129	Sequence 129, App
	29	15	53.6	31	1	US-08-390-850-43	Sequence 43, App
	30	15	53.6	31	2	US-08-435-634-43	Sequence 43, App
	31	14.8	52.9	53	3	US-08-811-492-120	Sequence 120, App
	32	14.8	52.9	54	4	US-08-649-991-21	Sequence 21, App
	33	14.8	52.9	54	4	US-08-649-991-24	Sequence 24, App
	34	14.8	52.9	54	4	US-08-649-991-31	Sequence 31, App
	35	14.8	52.9	54	6	PCr-US94-09700-35	Sequence 35, App
	36	14.8	52.9	65	2	US-08-273-594-29	Sequence 29, App
	37	14.6	52.1	69	3	US-07-916-098A-26	Sequence 26, App
	38	14.6	52.1	69	3	US-07-916-098A-31	Sequence 31, App
	39	14.2	50.7	50	4	US-08-850-049-5	Sequence 5, App
	40	14.2	50.7	50	4	US-08-050-478-5	Sequence 9, App
	41	14.2	50.7	54	4	US-08-649-991-9	Sequence 19, App
	42	14.2	50.7	54	4	US-08-649-991-19	Sequence 19, App
	43	14.2	50.7	57	7	5514566-17	Sequence 69, App
	44	14.2	50.7	99	2	US-08-463-115-69	Sequence 69, App
	45	14.2	50.7	99	2	US-08-465-388-69	Sequence 69, App

ALIGNMENTS

RESULT 1
US-08-446-908-8/c
Sequence 8, Application US/08446908
Patent No. 5705149
GENERAL INFORMATION:
APPLICANT: Namen, Anthony E.
APPLICANT: Goodwin, Raymond G.
APPLICANT: Lupton, Stephen D.
APPLICANT: Mochizuki, Diane Y.
TITLE OF INVENTION: Interleukin-7 and Antibodies Reactive
TITLE OF INVENTION: Therewith
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.1
SOFTWARE: Microsoft Word, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,908
FILING DATE: 22-May-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/231,205
FILING DATE: 21-Apr-1994
APPLICATION NUMBER: US 07/957,649
FILING DATE: 06-Oct-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,438
FILING DATE: 13-Apr-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/255,209
FILING DATE: 07-Oct-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/113,566
FILING DATE: 26-Oct-1987
ATTORNEY/AGENT INFORMATION:
NAME: Seese, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2104-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 8:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 19..60
; US-08-446-908-8

Query Match          57.1%; Score 16; DB 2; Length 79;
Best Local Similarity 79.2%; Pred. No. 50;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ctaccgtgattgctggcgc 27
    | |||| | |||| | |||| | ||||
Db 57 CCTACCTTCGATTTCCTTGTCTATC 34

RESULT 2
US-08-231-205A-8/c
; Sequence 8, Application US/08231205A
; Patent No. 5714585
; GENERAL INFORMATION:
; APPLICANT: Namen, Anthony E.
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Lupton, Stephen D.
; APPLICANT: Mochizuki, Diane Y.
; TITLE OF INVENTION: Interleukin-7 and Antibodies Reactive
; TITLE OF INVENTION: Therewith
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple 7.1
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,205A
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/957,649
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,438
; FILING DATE: 13-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/255,209
; FILING DATE: 07-OCT-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/113,566
; FILING DATE: 26-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2104-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```

; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 19..60
; US-08-231-205A-8

Query Match          57.1%; Score 16; DB 2; Length 79;
Best Local Similarity 79.2%; Pred. No. 50;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ctaccgtgattgctggcgc 27
    | |||| | |||| | |||| | ||||
Db 57 CCTACCTTCGATTTCCTTGTCTATC 34

RESULT 3
US-08-871-161-8/c
; Sequence 8, Application US/08871161
; Patent No. 5965122
; GENERAL INFORMATION:
; APPLICANT: Namen, Anthony E.
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Lupton, Stephen D.
; APPLICANT: Mochizuki, Diane Y.
; TITLE OF INVENTION: Interleukin-7 and Antibodies Reactive
; TITLE OF INVENTION: Therewith
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple 7.1
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,161
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,908
; FILING DATE: 22-MAY-1995
; APPLICATION NUMBER: US 08/231,205
; FILING DATE: 21-APR-1994
; APPLICATION NUMBER: US 07/957,649
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,438
; FILING DATE: 13-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/255,209
; FILING DATE: 07-OCT-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/113,566
; FILING DATE: 26-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2104-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```


CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-12

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttctaccgttattgcttgacatc 28
||||| ||| ||||| | |||
DB 47 TTCCTACTTTGTTTGCTCTCTCT 21

RESULT 7
US-08-649-991-13/c
Sequence 13, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Narva, Remy
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
MATERNOFETAL TRANSMISSION OF HIV-1
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995

ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-13

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ttctaccgttattgcttgacatc 28
||||| ||| ||||| | |||
DB 47 TTCCTACTTTGTTTGCTCTCTCT 21

RESULT 8
US-08-649-991-14/c
Sequence 14, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Narva, Remy
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
MATERNOFETAL TRANSMISSION OF HIV-1
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-14

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ttctacgttgattgctgcatc 28
||||| ||| ||| ||| |||
DB 47 ttcttacttttggttgctctctct 21

RESULT 9
US-08-649-991-15/c
; Sequence 15, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-649-991-15

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ttctacgttgattgctgcatc 28
||||| ||| ||| ||| |||
DB 47 ttcttacttttggttgctctctct 21

RESULT 10
US-08-649-991-16/c
; Sequence 16, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy

APPLICANT: Roques, Pierre
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ttctacgttgattgctgcatc 28
||||| ||| ||| ||| |||
DB 47 ttcttacttttggttgctctctct 21

RESULT 11
US-08-649-991-17/c
; Sequence 17, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7176
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-17

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ttctacgtgattgctgcatcct 28
||||| ||| ||||| | |||
Db 47 ttctactttgttcttgcctcctcct 21

RESULT 12
US-08-649-991-18/c
Sequence 18, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Naiva, Remy
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7176
TELEFAX: 202-467-7176

INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-18

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ttctacgtgattgctgcatcct 28
||||| ||| ||||| | |||
Db 47 ttctactttgttcttgcctcctcct 21

RESULT 13
US-08-649-991-20/c
Sequence 20, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Naiva, Remy
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7176
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-20

Query Match 56.4%; Score 15.8; DB 4; Length 54;
Best Local Similarity 74.1%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ttctacgtgattgctgcatcct 28
||||| ||| ||||| | |||

Db 47 TTTCTACTTTTGTGCTCTCTCT 21

RESULT 14

US-08-649-991-22/c

Sequence 22, Application US/08649991

Patent No. 5919462

GENERAL INFORMATION:

APPLICANT: Naiva, Remy

TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE

TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR

TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF

NUMBER OF SEQUENCES: 130

CORRESPONDENCE ADDRESSES:

ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP

STREET: 1800 M Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036-5869

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/649,991

FILING DATE: 17-MAY-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 9505914

FILING DATE: 18-MAY-1995

ATTORNEY/AGENT INFORMATION:

NAME: Adler, Reid G.

REGISTRATION NUMBER: 30,988

REFERENCE/DOCKET NUMBER: ORES-5003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-467-7000

TELEFAX: 202-467-7176

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 54 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-649-991-22

Query Match 56.4%; Score 15.8; DB 4; Length 54;

Best Local Similarity 74.1%; Pred. No. 57;

Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ttctacgctgattgctgcctc 28

Db 47 TTTCTACTTTTGTGCTCTCTCT 21

RESULT 15

US-08-649-991-23/c

Sequence 23, Application US/08649991

Patent No. 5919462

GENERAL INFORMATION:

APPLICANT: Naiva, Remy

TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE

TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR

TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF

NUMBER OF SEQUENCES: 130

CORRESPONDENCE ADDRESSES:

ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP

STREET: 1800 M Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036-5869

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/649,991

FILING DATE: 17-MAY-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 9505914

FILING DATE: 18-MAY-1995

ATTORNEY/AGENT INFORMATION:

NAME: Adler, Reid G.

REGISTRATION NUMBER: 30,988

REFERENCE/DOCKET NUMBER: ORES-5003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-467-7000

TELEFAX: 202-467-7176

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 54 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-649-991-23

Query Match 56.4%; Score 15.8; DB 4; Length 54;

Best Local Similarity 74.1%; Pred. No. 57;

Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ttctacgctgattgctgcctc 28

Db 47 TTTCTACTTTTGTGCTCTCTCT 21

Search completed: June 4, 2000, 16:09:42

Job time: 28061 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:33 ; Search time 1236.38 Seconds

(without alignments)
-22.031 Million cell updates/sec

Title: US-09-164-714-11

Perfect score: 28
Sequence: 1 ctgaccaatttggcaccacaacatagag 28

Scoring table: IDENTITY_NUC
Gapop 10.0 / Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_hc91:*
33: gb_hc92:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_hc93:*
42: gb_hc94:*
43: gb_hc95:*
44: gb_hc96:*

45: gb_hc97:*
46: em_hc91:*
47: em_hc92:*
48: em_hc93:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_hc98:*
53: gb_hc99:*
54: gb_hc10:*
55: gb_hc911:*
56: gb_hc912:*
57: gb_hc913:*
58: gb_hc914:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	16	57.1	39	5	AR012248	AR012248 Sequence
2	16	57.1	39	5	I14946	I14946 Sequence 32
3	16	57.1	39	5	I73666	I73666 Sequence 32
4	15	53.6	34	5	A39689	A39689 Sequence 2
5	15	53.6	34	5	A48961	A48961 Sequence 8
6	14.8	52.9	75	40	AF064903	AF064903 Homo sapi
7	14.8	52.9	92	10	HUMPHOS15	M59733 Human muscl
8	14.2	50.7	53	4	FR08001	AJ008001 Fugu rubr
9	14.2	50.7	60	8	AF101150	AF101150 Brassica
10	14.2	50.7	72	5	AR040932	AR040932 Sequence
11	14.2	50.7	72	5	AR051839	AR051839 Sequence
12	14.2	50.7	72	5	I36229	I36229 Sequence 31
13	14.2	50.7	90	9	GORINSORFA	M69096 Gorilla end
14	14.2	50.7	99	34	DD1DDKE	M59748 D.discoideu
15	14	50.0	27	5	E14994	E14994 PCR primer
16	14	50.0	36	5	A78745	A78745 Sequence 6
17	14	50.0	36	5	AR014718	AR014718 Sequence
18	13.8	49.3	35	5	AR055776	AR055776 Sequence
19	13.8	49.3	35	5	AR055777	AR055777 Sequence
20	13.8	49.3	51	12	U92178	U92178 Mus musculu
21	13.6	48.6	26	5	AR014478	AR014478 Sequence
22	13.6	48.6	59	10	S59972	S59972 TAL1d1-prot
23	13.6	48.6	60	10	S72312	S72312 phosphofuc
24	13.6	48.6	81	3	PIGBROC	M25036 Hog progest
25	13.6	48.6	99	12	MUSLRP03	L13679 Mouse leuko
26	13.4	47.9	44	5	AR048165	AR048165 Sequence
27	13.4	47.9	77	13	R1C1254A	D25437 Rice genom1
28	13.4	47.9	87	10	S71607	S71607 pol (clone
29	13.2	47.1	23	5	AR063157	AR063157 Sequence
30	13.2	47.1	29	5	AR003150	AR003150 Sequence
31	13.2	47.1	29	5	AR003243	AR003243 Sequence
32	13.2	47.1	29	5	AR009108	AR009108 Sequence
33	13.2	47.1	29	5	AR011395	AR011395 Sequence
34	13.2	47.1	29	5	AR028081	AR028081 Sequence
35	13.2	47.1	29	5	AR052797	AR052797 Sequence
36	13.2	47.1	29	5	I18903	I18903 Sequence 26
37	13.2	47.1	43	5	A05133	A05133 Oligonucleo
38	13.2	47.1	50	9	HPBHUMIN2	K01660 Hepatitis B
39	13.2	47.1	72	5	I00803	I00803 Sequence 7
40	13.2	47.1	77	5	A00281	A00281 Modified fr
41	13.2	47.1	77	5	A16012	A16012 HSA oligonu
42	13.2	47.1	93	12	M676712	U86712 Mus musculu
43	13	46.4	21	5	AR024394	AR024394 Sequence
44	13	46.4	21	5	I71186	I71186 Sequence 38
45	13	46.4	21	5	I83745	I83745 Sequence 38

ALIGNMENTS

RESULT 1
AR012248 AR012248 39 bp DNA PAT 04-DEC-1998
LOCUS Sequence 32 from patent US 5763254.
DEFINITION AR012248
ACCESSION AR012248
VERSION AR012248.1 GI:3970238
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 39)
TITLE Enzyme capable of degrading cellulose or hemicellulose
JOURNAL Patent: US 5763254-A 32 09-JUN-1998;
FEATURES Location/Qualifiers
source 1..39
BASE COUNT 9 a 16 c 8 g 6 t
ORIGIN

Query Match 57.1%; Score 16; DB 5; Length 39;
Best Local Similarity 79.2%; Pred. NO. 2.5e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 gaccatttggcaccacatag 27
||||| ||| | ||| |
Db 11 GACCAATTCTGGCCACACATTGG 34

RESULT 2
114946 39 bp DNA PAT 02-APR-1996
LOCUS Sequence 32 from patent US 5457046.
DEFINITION 114946
ACCESSION 114946
VERSION 114946.1 GI:1249854
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 39)
TITLE Woldike,H.F., Hagen,F., Hjort,C. and Hastrup,S.
JOURNAL Enzyme capable of degrading cellulose or hemicellulose
FEATURES Patent: US 5457046-A 32 10-OCT-1995;
source Location/Qualifiers
1..39
/organism="unknown"
BASE COUNT 9 a 16 c 8 g 6 t
ORIGIN

Query Match 57.1%; Score 16; DB 5; Length 39;
Best Local Similarity 79.2%; Pred. NO. 2.5e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 gaccatttggcaccacatag 27
||||| ||| | ||| |
Db 11 GACCAATTCTGGCCACACATTGG 34

RESULT 3
173666 39 bp DNA PAT 03-APR-1998
LOCUS Sequence 32 from patent US 5686593.
DEFINITION 173666
ACCESSION 173666
VERSION 173666.1 GI:3009807
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 39)
TITLE Woldike,H.Fabritius, Hagen,F., Hjort,C.Malland and Hastrup,S.

TITLE Enzyme capable of degrading cellulose or hemicellulose
JOURNAL Patent: US 5686593-A 32 11-NOV-1997;
FEATURES Location/Qualifiers
source 1..39
/organism="unknown"
BASE COUNT 9 a 16 c 8 g 6 t
ORIGIN

Query Match 57.1%; Score 16; DB 5; Length 39;
Best Local Similarity 79.2%; Pred. NO. 2.5e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 gaccatttggcaccacatag 27
||||| ||| | ||| |
Db 11 GACCAATTCTGGCCACACATTGG 34

RESULT 4
A39689/c 34 bp DNA PAT 05-MAR-1997
LOCUS Sequence 2 from Patent WO9418327.
DEFINITION A39689
ACCESSION A39689
VERSION A39689.1 GI:2295950
KEYWORDS
SOURCE
ORGANISM unidentified.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 34)
TITLE Burnham,M.K., Chopra,I., Critchley,I.A. and Knowles,D.J.
JOURNAL FIBRONECTIN BINDING PROTEIN; MONOCLONAL ANTIBODY AND THEIR USE IN PREVENTING BACTERIAL ADHESION
COMMENT Patent: WO 9418327-A 2 18-AUG-1994;
SMITHKLINE BEECHAM PLC (GB)
Other publication AU 5975994 940829
Other publication CN 1119026 960320
Other publication CN 2155413 940818
Other publication ZA 9400728 941107
Other publication JP 8506810F 960723.
source Location/Qualifiers
1..34
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 5 a 6 c 14 g 9 t
ORIGIN

Query Match 53.6%; Score 15; DB 5; Length 34;
Best Local Similarity 78.3%; Pred. NO. 7.2e+03;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 6 ccaatttggcaccacatagg 28
||||| ||| ||| ||| |||
Db 29 CCAATCTGCGCCACACGTAAGG 7

RESULT 5
A48961/c 34 bp DNA PAT 07-MAR-1997
LOCUS Sequence 8 from Patent WO9604380.
DEFINITION A48961
ACCESSION A48961
VERSION A48961.1 GI:2302599
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 34)
TITLE Critchley,I.A., Dodd,I., Barnett,P. and Mossakowska,D.E.
JOURNAL D2D3 POLYPEPTIDE FROM STAPHYLOCOCCUS AUREUS AND USES THEREOF
COMMENT Patent: WO 9604380-A 8 15-FEB-1996;
SMITHKLINE BEECHAM PLC (GB)
Other publication AU 3223895 960304.
source Location/Qualifiers

```

source
1. .34
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT      5 a      6 c      14 g      9 t
ORIGIN

Query Match      53.6%; Score 15; DB 5; Length 34;
Best Local Similarity 78.3%; Pred. No. 7.2e+03;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY      6 ccaatttgcacacacatag 28
      ||||| ||||| ||||| |||||
Db      29 CCAATGTCACCACTAAG 7

RESULT      6
AF064903/c      75 bp      mRNA      PRI      03-MAR-1999
LOCUS      AF064903
DEFINITION      Homo sapiens clone 12-1 Immunoglobulin heavy chain alpha VDJ region
ACCESSION      AF064903
VERSION      AF064903.1 GI:4321896
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 75)
AUTHORS      Rowley,A.H., Eckerley,C.A., Shulman,S.T. and Baker,S.C.
TITLE      Evidence of an antigen-driven IgA immune response in the vascular
wall in acute Kawasaki Syndrome
JOURNAL      Unpublished
AUTHORS      2 (bases 1 to 75)
AUTHORS      Rowley,A.H., Eckerley,C.A., Shulman,S.T. and Baker,S.C.
TITLE      Direct Submission
JOURNAL      Submitted (14-MAY-1998) Pediatrics, Loyola University, 2160 S First
Avenue, Maywood, IL 60153, USA
FEATURES
SOURCE      Location/Qualifiers
1. .75
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="12-1"
/lisse="type="vascular"
/notes="from aorto-iliac junction of a 10-year-old child
with fatal acute Kawasaki Syndrome"
<1. .>75
/notes="Includes IGHV3 gene"
/codon_start=1
/product="Immunoglobulin heavy chain alpha VDJ region"
/protein_id="AADI5867.1"
/db_xref="GI:4321897"
/translation="AVYYCAKNROVQPFHYHYMDVWG"

BASE COUNT      19 a      22 c      15 g      19 t
ORIGIN

Query Match      52.9%; Score 14.8; DB 40; Length 75;
Best Local Similarity 73.1%; Pred. No. 9.4e+03;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY      1 cttgaccatttgcacacacatag 26
      ||||| ||||| ||||| |||||
Db      31 CTTGACGATTTTCGACACGTAATAG 6

RESULT      7
HUMMPHOS15/c      92 bp      DNA      PRI      05-DEC-1995
LOCUS      HUMMPHOS15
DEFINITION      Human muscle-type phosphofructokinase (PFK-M) gene, exon 16.
ACCESSION      M59733
VERSION      M59733.1 GI:188647
KEYWORDS      muscle-type phosphofructokinase.

```

```

SEGMENT      15 of 23
SOURCE      Homo sapiens DNA.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 92)
AUTHORS      Yamasaki,T., Nakajima,H., Kono,N., Hotta,K., Yamada,K., Imai,E.,
Kunajima,M., Noguchi,T., Tanaka,T. and Tani,T.
TITLE      Structure of the entire human muscle phosphofructokinase-encoding
gene: a two-promoter system
JOURNAL      Gene 104 (2), 277-282 (1991)
MEDLINE      92009225
FEATURES
SOURCE      Location/Qualifiers
1. .92
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1cen-q32"
16. .86
/gene="PFK-M"
/notes="600-120-277"
/number=16

BASE COUNT      23 a      17 c      32 g      20 t
ORIGIN

Query Match      52.9%; Score 14.8; DB 10; Length 92;
Best Local Similarity 73.1%; Pred. No. 9.5e+03;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY      1 cttgaccatttgcacacacatag 26
      ||||| ||||| ||||| |||||
Db      61 CTTGCGCATGTCACGCCCACTAAG 36

RESULT      8
FR08001      53 bp      DNA      VRT      18-JAN-1999
LOCUS      FR08001
DEFINITION      Fugu rubripes Eya4 gene, exon 4.
ACCESSION      AJ008001
VERSION      AJ008001.1 GI:4164315
KEYWORDS      EYA4 gene.
SOURCE      Fugu rubripes.
ORGANISM      Fugu rubripes.
Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
Tetraodontiformes; Tetraodontidae; Tetraodontidae; Fugu.
REFERENCE      1 (bases 1 to 53)
AUTHORS      Hanson,I.M.
TITLE      Direct Submission
JOURNAL      Submitted (16-JUL-1998) Hanson I.M., Cell Genetics Section, MRC
Human Genetics Unit, Western General Hospital, Crewe Road,
Edinburgh, EH4 2XU, UK
2 (bases 1 to 53)
Borsani,G., deGrandi,A., Ballabio,A., Bulfone,A., Bernard,L.,
Banfi,S., Gattuso,C., Mariani,M., Dixon,M., Donati,D., Metcalfe,K.,
Winter,R., Robertson,M., Axton,R., Brown,A., van Heyningen,V. and
Hanson,I.
TITLE      EYA4, a novel vertebrate gene related to drosophila eyes absent
JOURNAL      Hum. Mol. Genet. 8 (1), 11-23 (1999)
MEDLINE      99105912
FEATURES
SOURCE      Location/Qualifiers
1. .53
/organism="Fugu rubripes"
/db_xref="taxon:31033"
/clone="cosmid 032C09"
1. .53
/gene="Eya4"
1. .>53
/gene="Eya4"
/number=4
/usedin=AJ007998:Eya4_CDS

BASE COUNT      14 a      11 c      12 g      16 t
ORIGIN

```


RESULT 13
GORINSORA 90 bp DNA PRI 27-APR-1993
LOCUS GORILLA endogenous retroviral DNA.
DEFINITION M69096
ACCESSION M69096
VERSION M69096.1 GI:177060
KEYWORDS
SOURCE
ORGANISM
Gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Gorilla.
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
1 (bases 1 to 90)
Shih,A., Coutavas,E.E. and Rush,M.G.
Evolutionary implications of primate endogenous retroviruses
Virology 182, 495-502 (1991)
91220699
FEATURES
source
Location/Qualifiers
1..90
/organism="Gorilla gorilla"
/db_xref="taxon:9593"
1..90
/partial
/note="ORF"
/codon_start=1
/protein_id="AA35469.1"
/db_xref="GI:177061"
/translation="FKNSPTLFDALHRLDIFRRHDPDLLQ"
BASE COUNT 19 a 31 c 15 g 25 t
ORIGIN

Query Match 50.7%; Score 14.2; DB 9; Length 90;
Best Local Similarity 84.2%; Pred. No. 1.8e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 cttgaccaatttgacc 19
||||| ||||| |||||
Db 45 CTTGACTGATTTGCAC 63

RESULT 14
DDIDDK 99 bp DNA INV 27-APR-1993
LOCUS D.discoidium protein kinase 5 gene, partial cds.
DEFINITION M59748
ACCESSION M59748
VERSION M59748.1 GI:167725
KEYWORDS
SOURCE
ORGANISM
Dictyostelium discoidium (strain AX-3) DNA.
Dictyostelium discoidium
Eukaryota; Dictyostelida; Dictyostelium.
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
1 (bases 1 to 99)
Haribabu,B. and Dotlin,R.P.
Identification of a protein kinase multigene family of a
Dictyostelium discoidium: Molecular cloning and expression of a
cDNA encoding a developmentally regulated protein kinase
Proc. Natl. Acad. Sci. U.S.A. 88, 1115-1119 (1991)
91142122
FEATURES
source
Location/Qualifiers
1..99
/organism="Dictyostelium discoidium"
/strain="AX-3"
/db_xref="taxon:44689"
1..99
/codon_start=1
/product="protein kinase 5"
/protein_id="AA33190.1"
/db_xref="GI:167726"
/translation="NILLTSDGHIVLTPFGISKGLVSDNDRTATFC"
BASE COUNT 30 a 18 c 24 g 27 t
ORIGIN

Query Match 50.7%; Score 14.2; DB 34; Length 99;

Best Local Similarity 70.4%; Pred. No. 1.8e+04;
Matches 19; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 ttgaccaatttgaccacacacatagg 28
||||| ||||| | | | | |
Db 34 TTGACCGATTTTGTGTAAGCAAGAG 60

RESULT 15
E14994/c 27 bp DNA PAT 28-JUL-1999
LOCUS PCR primer to mutagenize equine cDNA encoding chorionic
DEFINITION gonadotropin alpha subunit at the 58th Aa to Gln.
ACCESSION E14994
VERSION E14994.1 GI:5709677
KEYWORDS JP 1998036285-A/8.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
1 (bases 1 to 27)
Ogawa,T., Shioda,K., Bin,K. and Ikemi,M.
AUTHORS NEW GONADOTROPIC HORMONE AND ITS PRODUCTION
TITLE Patent: JP 1998036285-A 10-FEB-1998;
JOURNAL DENKI KAGAKU KOGYO KK
COMMENT
OS None
OC Artificial sequences.
PN JP 1998036285-A/8
PD 10-FEB-1998
PF 23-JUL-1996 JP 1996193232
PI OGAWA TOMOYA, SHIODA KUNIO, BIN KANSHIYOKU, IKEMI MASAHA PC
A61K38/24,C07K14/59,C12N5/10,C12N15/09,C12P21/02,C12P21/02, PC
C12R1:91);
CC strandness: Single;
FH Key
FH topology: Linear;
FH Key
FT source
FT 1..27
Location/Qualifiers
1..27
/organism="Artificial sequences"
/db_xref="taxon:32644"
BASE COUNT 8 a 8 c 5 g 6 t
ORIGIN

Query Match 50.0%; Score 14; DB 5; Length 27;
Best Local Similarity 77.3%; Pred. No. 2.1e+04;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 tgaccaatttgaccacacat 24
||| | ||||| |||||
Db 22 TGATCTGCTTTGGACCAACAT 1

Search completed: June 4, 2000, 16:05:34
Job time: 27886 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:12 ; Search time 4521.53 Seconds
(without alignments)
16.136 Million cell updates/sec

Title: us-09-164-714-12
Perfect score: 18
Sequence: 1 ggcgacaacacgcctag 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *
44: gb_est25: *

45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_gss1: *
83: gb_gss2: *
84: gb_gss3: *
85: gb_gss4: *
86: em_gss1: *
87: em_gss2: *
88: em_gss3: *
89: em_gss4: *
90: gb_gss5: *
91: gb_gss6: *
92: gb_gss7: *
93: gb_gss8: *
94: gb_gss9: *
95: em_gss5: *
96: em_gss6: *
97: em_gss7: *
98: em_gss8: *
99: em_gss9: *
100: em_gss10: *
101: em_gss11: *
102: gb_gss10: *
103: gb_gss11: *
104: em_gss12: *
105: gb_gss13: *
106: gb_gss14: *
107: gb_gss15: *
108: gb_gss16: *
109: gb_gss17: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
c 1	13.4	74.4	85	25	U44241	U44241 ENU44241 AS
c 2	13.2	73.3	87	40	AA971367	AA971367 op80c11.s
c 3	12.8	71.1	46	36	AA607492	AA607492 w61h02.r
c 4	12.8	71.1	97	29	AA151757	AA151757 z029f01.r
c 5	12.8	71.1	97	43	AA181902	AA181902 ud73e07.x
c 6	12.4	68.9	53	27	C01203	C01203 HUMGS000790
c 7	12.4	68.9	60	40	AA938614	AA938614 nw91h09.s
c 8	12.4	68.9	67	47	AI472550	AI472550 l178a02.x
c 9	12.4	68.9	74	30	AA218675	AA218675 zq14f10.s
c 10	12.4	68.9	86	34	AA515609	AA515609 ng70a03.s
c 11	12.2	67.8	55	43	AI220035	AI220035 q978h07.x
c 12	12.2	67.8	79	43	AI214486	AI214486 q969d07.x
c 13	12.2	67.8	88	51	AI744149	AI744149 wc36h02.x
c 14	12.2	67.8	95	30	AA244496	AA244496 nc07910.s
c 15	12.2	67.8	99	94	AQ025006	AQ025006 EP(2)0969
c 16	12.2	67.8	88	44	AI310822	AI310822 q093h04.x
c 17	11.8	65.6	52	40	AA934866	AA934866 op47a10.s
c 18	11.8	65.6	61	39	AA894395	AA894395 of85903.s
c 19	11.8	65.6	66	94	AQ073497	AQ073497 EP(2)2329
c 20	11.8	65.6	74	46	AI445897	AI445897 t106905.x
c 21	11.8	65.6	76	38	AA789177	AA789177 aq58c03.s
c 22	11.8	65.6	82	61	AI865432	AI865432 wk11h10.x
c 23	11.8	65.6	100	30	AA247764	AA247764 hfe0467.s
c 24	11.8	65.6	100	43	AI204724	AI204724 ZF-EST40
c 25	11.6	64.4	88	39	AA864661	AA864661 oh37c11.s
c 26	11.6	64.4	91	43	AI181540	AI181540 uc59g06.r
c 27	11.6	64.4	94	29	AA144170	AA144170 mg54g10.r
c 28	11.6	64.4	97	44	AI323709	AI323709 mq37c06.x
c 29	11.6	64.4	100	81	AA453587	AA453587 zeh10164
c 30	11.4	63.3	34	38	AA878903	AA878903 vl79f06.r
c 31	11.4	63.3	63	44	AI286217	AI286217 q102e01.x
c 32	11.4	63.3	65	37	AA733449	AA733449 vl73h08.r
c 33	11.4	63.3	82	34	AA533454	AA533454 nj68c05.s
c 34	11.2	62.2	37	34	AA466917	AA466917 vl10904.r
c 35	11.2	62.2	37	43	AI188273	AI188273 qd11g10.x
c 36	11.2	62.2	44	42	AI153789	AI153789 vz90h06.r
c 37	11.2	62.2	44	94	AQ074027	AQ074027 EP(3)3471
c 38	11.2	62.2	49	28	C20877	C20877 HUMGS000494
c 39	11.2	62.2	50	28	AA108275	AA108275 EST00018.r
c 40	11.2	62.2	51	27	AA000409	AA000409 me76c01.r
c 41	11.2	62.2	51	81	AA424054	AA424054 sh60b02.y
c 42	11.2	62.2	58	34	AA495205	AA495205 fa04c02.r
c 43	11.2	62.2	67	29	AA157222	AA157222 z041d06.r
c 44	11.2	62.2	72	28	AA072178	AA072178 mme69f01.r
c 45	11.2	62.2	73	21	T91290	T91290 yd52c07.s1

ALIGNMENTS

RESULT 1
LOCUS U44241 85 bp mRNA EST 03-APR-1996
DEFINITION ENU44241 Aspergillus nidulans cleistothecium Emericella nidulans
ACCESSION U44241
VERSION U44241.1 GI:1244904
KEYWORDS EST.
SOURCE Emericella nidulans.
ORGANISM Emericella nidulans.
Eukaryota; Fungi; Ascomycota; Eurotiales; Trichocomaceae; Emericella.

REFERENCE 1 (bases 1 to 85)
AUTHORS Lee,D., Lee,S., Hwang,H., Kim,J. and Chae,K.
TITLE Quantitative analysis of gene expression in sexual structures of Aspergillus nidulans by sequencing of 3'-directed cDNA clones
JOURNAL FEBS Microbiol. Lett. 138 (1), 71-76 (1996)
MEDLINE 96236220
COMMENT On Apr 14, 1993 this sequence version replaced gi:838110.
Contact: Keon-Sang Chae
Chonbuk National University

Chonju, 561-756, S. Korea
Tel: +82-652-70-3340
Fax: +82-652-70-3345
Email: chae@chonbukns.chonbuk.ac.kr.
Location/Qualifiers

FEATURES

source

1. 85
/organism="Emericella nidulans"
/strain="FGSC4"
/db_xref="taxon:5072"
/clone="SE0504"
/clone_lib="Aspergillus nidulans cleistothecium"
/tissue_type="cleistothecium"
/cell_type="Hull cell"
/dev_stage="sexual"
/note="3'-directed cDNA clones; single-pass sequencing"

BASE COUNT 15 a 17 c 22 g 31 t
ORIGIN

Query Match 74.4%; Score 13.4; DB 25; Length 85;
Best Local Similarity 93.3%; Pred. No. 3.3e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 4 acaaacacgactag 18
Db 34 ACAAAACGACCTCG 20

RESULT 2

AA971367 87 bp mRNA EST 07-JUL-1998
LOCUS op80c11.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
DEFINITION IMAGE:1583156 3', mRNA sequence.
AA971367

ACCESSION AA971367.1 GI:3146657
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 87)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

On Jan 19, 1998 this sequence version replaced gi:2285140.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 892 Std Error: 0.00

Seq primer: -40m13 fwd. ER from Amersham
High quality sequence stop: 73.
Location/Qualifiers

FEATURES

source

1. 87
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1583156"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site.1: Not I; Site.2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung KBHL19W, testis NHT, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ACCESSION	alpha (mouse);, mRNA sequence.
AI181902	
AI181902.1	GI:3732540
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 97) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE	The WashU-HMHI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	On Jan 14, 1998 this sequence version replaced gi:1797360. Contact: Marra M/Mouse EST Project WashU-HMHI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:924872 Trace considered overall poor quality Seq primer: custom primer used High quality sequence stop: 1. location/Qualifiers 1..97 /organism="Mus musculus" /strain="C57BL" /db_xref="taxon:10090" /clone="IMAGE:1451556" /clone_lib="Sugano mouse liver mlia" /sex="female" /dev_stage="adult" /lab_host="DH10B" /note="Organ: liver; Vector: pME185-Fl3; Site:1: DraIII (CACTGTGTG); Site:2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor (TGTGGCCCTACTG), digested and cloned into distinct DraIII sites of the pME185-Fl3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGGCTCTTAAGCTGTG and 3' end primer CGACCTCGACCTCGACACA."
BASE COUNT	32 a 19 c 21 g 25 t
ORIGIN	
Query Match	71.1%; Score 12.8; DB 43; Length 97;
Best Local Similarity	87.5%; Pred. No. 6.9e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Oy 1 ggcacaaaccagcct 16	
Db 84 GTGACCAAAACGAGCT 69	
RESULT 6	
LOCUS C01203 53 bp mRNA EST 23-JUL-1996	
DEFINITION HUNGSO007903 Human adult (K.Okubo) Homo sapiens cDNA, mRNA sequence.	
ACCESSION C01203	
VERSION C01203.1 GI:1433433	

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS	Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE	1 (bases 1 to 53)
JOURNAL	Okubo, K.
COMMENT	BodyMap, human gene expression database Unpublished (1995) Contact: Okubo, K. Institute for Molecular and Cellular Biol Osaka University 1-3,Yamada-oka, Suita, Osaka Pref. 565, Japan Tel.: 06-877-5111(ex.3315) Email: kousaku@imcb.osaka-u.ac.jp Human Gene Signature, 3'-directed cDNA sequence. We are not submitting the same cDNA sequence redundantly to DBD since 1993. For the abundance information of clones with this sequence in this library and as well as in other 3'-directed libraries, see ' http://www.imcb.osaka-u.ac.jp/bodymap/ '. The sequences of the clones represented by this GS sequences is also found there.
FEATURES	Location/Qualifiers 1..53
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="Human adult (K.Okubo)" /dev_stage="adult"
BASE COUNT	10 a 15 c 16 g 12 t
ORIGIN	
Query Match	68.9%; Score 12.4; DB 27; Length 53;
Best Local Similarity	92.9%; Pred. No.1e+04;
Matches 13: Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Oy	5 caaacacgacctag 18
Db	32 CAAAACGAGCCCG 19
RESULT 7	
LOCUS	AA938614 60 bp mRNA EST 30-APR-1998
DEFINITION	nw11h09.s1 NCI-CGAP.Pr12 Homo sapiens CDNA clone IMAGE:1254017
ACCESSION	similar to gb:U01120 GUDCOSE-6-PHOSPHATASE (HUMAN); mRNA sequence.
VERSION	AA938614
KEYWORDS	AA938614.1 GI:3096642
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE	1 (bases 1 to 60) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) On Jan 19, 1998 this sequence version replaced gi:2152774. Contact: Robert Strausberg, Ph.D. Tel.: (301) 496-1550 Email: Robert_Strausberg@nih.gov Tissue Procurement: W. Douglas Figg, Ph.D., Paul H. Duray, M.D., Rodrigo F. Chuquai, M.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: David B. Krizman, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/JLML at: www-bio.llnl.gov/bbrp/image/image.html
JOURNAL	
COMMENT	Trace considered overall poor quality Seq primer: -40m13 fwd. ET from Amersham High quality sequence stop:.1.

FEATURES
source
Location/Qualifiers
1. .60
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1254017"
/clone_lib="NCI_CGAP_Pri12"
/sex="male"
/tissue_type="metastatic prostate bone lesion"
/lab_host="DH10B"
/note="Vector: PAMPI0; mRNA made from metastatic prostate lesion of the bone, cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Library made by D. Krizman, NIH."

BASE COUNT
ORIGIN
9 a 13 c 13 g 25 t

Query Match
Best Local Similarity 68.9%; Score 12.4; DB 40; Length 60;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 caaaccagcctag 18
|||||

DB 48 CAAGACCAAGCCTAG 35

RESULT 8
AI472550/c
LOCUS t178a02.x1 NCI_CGAP_Col4 Homo sapiens cDNA clone IMAGE:2153162 3',
DEFINITION mRNA sequence.
ACCESSION AI472550
VERSION AI472550.1 GI:4325595
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 67)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jun 5, 1998 this sequence version replaced gi:3187297.
CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/imap/image.html

FEATURES
source
Insert Length: 1410 Std Error: 0.00
Seg primer: -40UP from gibco
High quality sequence stop: 61.
Location/Qualifiers
1. .67
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2153162"
/clone_lib="NCI_CGAP_Col4"
/tissue_type="moderately-differentiated adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: PCMV-SPORE6; Site: 1: SalI;
Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.7 kb. Life Technologies catalog #: 11531-019"

BASE COUNT
ORIGIN
11 a 16 c 15 g 25 t

ORIGIN
Query Match
Best Local Similarity 68.9%; Score 12.4; DB 47; Length 67;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 caaaccagcctag 18
|||||

DB 41 CAAGACCAAGCCTAG 28

RESULT 9
AA218675/c
LOCUS zq14f10.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
DEFINITION IMAGE:629707 3' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION AA218675
VERSION AA218675.1 GI:1832759
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 74)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,
Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rife, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J.,
Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
TITLES
JOURNAL
MEDLINE
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393955.
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41m3 fwd. ET from Amersham
High quality sequence stop: 51.
Location/Qualifiers
1. .74
/organism="Homo sapiens"
/db_xref="GDB:5179779"
/db_xref="taxon:9606"
/clone="IMAGE:629707"
/clone_lib="Stratagene fetal retina 937202"
/sex="mixed"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site: 1: EcoRI; Site: 2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled
retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
Vector: -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3'
adaptor sequence: 5' CTCGAGCTTTTTCCTTTTTCCTTTT 3' "

BASE COUNT
ORIGIN
15 a 18 c 16 g 24 t 1 others

Query Match
Best Local Similarity 68.9%; Score 12.4; DB 30; Length 74;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 caaaccagcctag 18
|||||

DB 46 CAAACCAAGCCTAG 33

RESULT 10
AA515609
LOCUS
DEFINITION
AA515609 86 bp mRNA EST 19-AUG-1997
ng70a03.s1 NCI-CGAP_Lip2 Homo sapiens cDNA clone IMAGE:940108
similar to contigs A19 repetitive element, contains element THR
repetitive element ; mRNA sequence.
ACCESSION
AA515609
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS
TITLE
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 9, 1995 this sequence version replaced gi:802649.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Insert Length: 591 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 76.
FEATURES
Source
Location/Qualifiers
1..86
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:940108"
/clone_lib="NCI-CGAP_Lip2"
/tissue_type="liposarcoma"
/lab_host="DH10B"
/note="Vector: pAMP10; mRNA made from liposarcoma, CDNA
made by oligo-dT priming. Non-directionally cloned.
Size-selected on agarose gel, average insert size 600 bp.
Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."
BASE COUNT 28 a 21 c 26 g 11 t
ORIGIN

Query Match 68.9%; Score 12.4; DB 34; Length 86;
Best Local Similarity 92.9%; Pred. No. 1.1e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 caaacacgacctag 18
|||||
Db 41 CAAACGACGCTCG 54

RESULT 11
A1220035
LOCUS
DEFINITION
A1220035 55 bp mRNA EST 30-NOV-1998
qg78h07.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:1841341 3' similar to SW:ZNI6_HUMAN P17020 ZINC FINGER
PROTEIN 16 ; mRNA sequence.
ACCESSION
A1220035
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

REFERENCE
AUTHORS
TITLE
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1198 Std Error: 0.00
Seq primer: -400P from Glibco
High quality sequence stop: 1.
FEATURES
Source
Location/Qualifiers
1..55
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1841341"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT7n3D-Pac (Pharmacia) with
a modified polylinker; Site:1: Not I; Site:2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CGAP (CDB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries.
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bernaldo."
BASE COUNT 10 a 21 c 17 g 7 t
ORIGIN

Query Match 67.8%; Score 12.2; DB 43; Length 55;
Best Local Similarity 82.4%; Pred. No. 1.3e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ggcacaaacgacctc 17
|||
Db 30 GCTGCAAAACCAACCTA 46

RESULT 12
A1214486
LOCUS
DEFINITION
A1214486 79 bp mRNA EST 21-OCT-1998
qg59d07.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:1840429 3' similar to TR:O14978 O14978 ZINC FINGER PROTEIN
FPM315 ; mRNA sequence.
ACCESSION
A1214486
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS
TITLE
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 17, 1998 this sequence version replaced gi:2043831.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Seq primer: -400P from Glibco

FEATURES High quality sequence stop: 1.

SOURCE

Location/Qualifiers

1..79

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1840429"

/clone_lib="Soares_NFL_T_GBC_S1"

/lab_host="PH10B"

/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with

a modified polylinker; Site_1: Not I; Site_2: Eco RI;

Equal amounts of plasmid DNA from three normalized

libraries (Fetal lung NbHL19w, testis NHT, and B-cell

NCI CGAP GCBI) were mixed, and ss circles were made in

vitro. Following HAP purification, this DNA was used as

tracer in a subtractive hybridization reaction. The driver

was PCR-amplified cDNAs from pools of 5,000 clones made

from the same 3 libraries. The pools consisted of

I.M.A.G.E. clones 297480-302087, 682632-687239,

726408-728711, and 729096-731399. Subtraction by Bento

Soares and M. Fatima Bonaldo.

BASE COUNT

16 a 33 c 15 g 15 t

ORIGIN

Query Match 67.8%; Score 12.2; DB 43; Length 79;
 Best Local Similarity 82.4%; Pred. No. 1.4e+04;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gcgacaaaccagccta 17

Db 30 GCCTCAAAACCACTA 46

RESULT 13

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AI744149 88 bp mRNA EST 17-DEC-1999
 W363b02.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2317299 3'
 similar to TR:Q13765 Q13765 NASCENT POLYPEPTIDE ASSOCIATED COMPLEX
 ALPHA SUBUNIT.; mRNA sequence.

AI744149

AI744149.1 GI:5112437

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 88)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On Dec 20, 1995 this sequence version replaced gi:1134761.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/dbfp/image/image.html

Trace considered overall poor quality

Insert length: 829 Std Error: 0.00

Seq primer: 400P from Gibco

High quality sequence stop: 1.

Location/Qualifiers

1..88

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2317299"

FEATURES

source

/clone_lib="NCI_CGAP_Pr28"

/sex="male"

/dev_stage="adult"

/lab_host="PH10B"

/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)

with a modified polylinker; Plasmid DNA from the

normalized library NCI CGAP Pr22 was prepared, and ss

circles were made in vitro. Following HAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (clonids

985608-986759, 1101192-1101959, and 1217928-1220615).

Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT

9 a 24 c 20 g 34 t 1 others

ORIGIN

Query Match 67.8%; Score 12.2; DB 51; Length 88;
 Best Local Similarity 82.4%; Pred. No. 1.4e+04;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 cgacaaaccagccta 18

Db 84 CCACACACACAGCCACAG 68

RESULT 14

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AA244496 95 bp mRNA EST 20-AUG-1997
 nc07g10.s1 NCI_CGAP_Pr1 Homo sapiens cDNA clone IMAGE:1007490, mRNA
 sequence.

AA244496

AA244496.1 GI:1875215

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 95)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On Apr 14, 1993 this sequence version replaced gi:503300.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,

M.D., Michael Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: David B. Kitzman, Ph.D.

DNA Sequencing by: Genome Systems Inc., Greg Lennon, Ph.D.

clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/dbfp/image/image.html

Seq primer: -41ml3 fwd. ET from Amersham

High quality sequence stop: 82.

Location/Qualifiers

1..95

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1007490"

/clone_lib="NCI_CGAP_Pr1"

/sex="Male"

/dev_stage="45 years old"

/lab_host="PH10B"

/note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st

strand cDNA was primed with oligo(dT)17 on 50 ng of

DNase-treated, total cellular RNA obtained from

5,000-10,000 microdissected, histologically normal

prostate epithelial cells. Double-stranded cDNA was

ligated to EcoRI adaptors, 5 cycles of PCR applied to the

cdna with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDC-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

BASE COUNT 28 a 13 c 27 g 27 t

ORIGIN

Query Match 67.8%; Score 12.2; DB 30; Length 95;
Best Local Similarity 82.4%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 gcagacaacacagccta 17
|||
Db 71 GCACCAAAACCACTCTA 55

RESULT 15

A0025006

LOCUS

DEFINITION

A0025006 99 bp DNA GSS 14-OCT-1998
EP(2)0969 Drosophila melanogaster EP line Drosophila melanogaster
genomic Sequence recovered from 5' end of P element, genomic survey
sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster
fruit fly.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The BDGP gene disruption project: single EP element insertions
unpublished (1998)
Contact: Gerald Rubin
Berkeley Drosophila Genome Project
University of California, Berkeley
LSA Building, Berkeley, CA 94720-3200, USA
Fax: 5106439947
Email: gerry@fruitfly.berkeley.edu

Sequence recovery method was Inverse PCR.

Sequence orientation is forward strand relative to 5' end of P
element

The P element insertion position is base 92 in the 99 bases. This
insertion position refers to the first base of the 8 base target
recognition sequence.

Class: transposon-tagged.

FEATURES

SOURCE

Location/Qualifiers

1..99
/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone.lib="Drosophila melanogaster EP line"

/note="Inverse PCR was performed on Drosophila

melanogaster strains each of which contains a single EP

transposable element insertion. (The generation of these

insertion strains is described in North P, Szabo K, Bailey

A, Laverly T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes

V, Ansgore W, Cohen SM. 1998. Systematic gain-of-function

genetics in Drosophila. Development 6:1049-1057.) The

resultant fragment for each strain was directly sequenced

to determine the genomic sequence at the site of

insertion. Details of the protocols used can be found at

http://fruitfly.berkeley.edu/p-disrupt/Inverse_pcr.html."

23 a 23 c 20 g 33 t

Query Match 67.8%; Score 12.2; DB 94; Length 99;
Best Local Similarity 82.4%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 cgacaacacagcctag 18
|||
Db 37 CAACAAACCAACGAG 53

Search completed: June 4, 2000, 13:53:16
Job time: 20940 sec

BASE COUNT
ORIGIN

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:34 ; Search time 1236.38 Seconds
(without alignments)
-14.163 Million cell updates/sec

Title: US-09-164-714-12
Perfect score: 18
Sequence: 1 gcgcacaaccagcctag 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenBml : *
1: gb_ba1 : *
2: gb_ba2 : *
3: gb_om : *
4: gb_ov : *
5: gb_pat : *
6: gb_ph : *
7: gb_pl1 : *
8: gb_pl2 : *
9: gb_pr1 : *
10: gb_pr2 : *
11: gb_pr3 : *
12: gb_ro : *
13: gb_sts : *
14: gb_sy : *
15: gb_un : *
16: gb_v1 : *
17: em_fun : *
18: em_hum1 : *
19: em_hum2 : *
20: em_in : *
21: em_om : *
22: em_or : *
23: em_ov : *
24: em_pat : *
25: em_ph : *
26: em_pl : *
27: em_ro : *
28: em_sy : *
29: em_un : *
30: em_v1 : *
31: em_v2 : *
32: gb_htg1 : *
33: gb_htg2 : *
34: gb_in1 : *
35: gb_in2 : *
36: em_ba1 : *
37: em_ba2 : *
38: em_hum3 : *
39: em_hum4 : *
40: gb_pr4 : *
41: gb_htg3 : *
42: gb_htg4 : *
43: gb_htg5 : *
44: gb_htg6 : *

45: gb_htg7 : *
46: em_htg1 : *
47: em_htg2 : *
48: em_htg3 : *
49: em_hum5 : *
50: gb_pl3 : *
51: gb_pr5 : *
52: gb_htg8 : *
53: gb_htg9 : *
54: gb_htg10 : *
55: gb_htg11 : *
56: gb_htg12 : *
57: gb_htg13 : *
58: gb_htg14 : *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	13.8	76.7	25	5	E07840	E07840 Upstream ac
2	13.8	76.7	25	5	I72301	I72301 Sequence 10
3	13.8	76.7	25	5	I79557	I79557 Sequence 2
4	13.8	76.7	30	5	I72318	I72318 Sequence 27
5	13.8	76.7	31	5	I72311	I72311 Sequence 20
6	13.8	76.7	31	5	I72312	I72312 Sequence 21
7	13.8	76.7	31	5	I79569	I79569 Sequence 14
8	13.8	76.7	31	5	I79570	I79570 Sequence 15
9	13.8	76.7	50	5	I72317	I72317 Sequence 26
10	13.4	74.3	30	5	A36554	A36554 Sequence 8
11	13.2	73.3	78	12	M05ETNICO	M30932 Mouse DNA u
12	12.4	68.9	30	5	A36555	A36555 Sequence 9
13	12.2	67.8	36	5	A19278	A19278 oligonucleo
14	12.2	67.8	39	5	A36995	A36995 Sequence 46
15	12.2	67.8	39	5	A39203	A39203 Sequence 51
16	12.2	67.8	39	5	A39262	A39262 Sequence 46
17	12.2	67.8	49	5	A16464	A16464 oligonucleo
18	12.2	67.8	49	5	A16465	A16465 oligonucleo
19	12.2	67.8	50	5	I79573	I79573 Sequence 18
20	12.2	67.8	50	5	I79574	I79574 Sequence 19
21	12.2	67.8	55	5	A19300	A19300 oligonucleo
22	12.2	67.8	57	5	E12104	E12104 signal pept
23	12.2	67.8	57	5	I04131	I04131 Sequence 19
24	12.2	67.8	57	5	I05343	I05343 Sequence 1
25	12.2	67.8	57	5	I06744	I06744 Sequence 12
26	12.2	67.8	63	7	YSCSUC2B	M13627 Yeast (S.ce
27	12.2	67.8	64	5	A23341	A23341 Artificial
28	12.2	67.8	64	5	AR068519	AR068519 Sequence
29	12.2	67.8	68	14	SYNVSCTNV	M13346 Yeast (S.ce
30	12.2	67.8	69	5	AR038140	AR038140 Sequence
31	12.2	67.8	69	5	I79868	I79868 Sequence 8
32	12.2	67.8	70	5	A24425	A24425 SCFVLA link
33	12.2	67.8	75	5	E00281	E00281 DNA coding
34	12.2	67.8	75	5	E00283	E00283 DNA coding
35	12.2	67.8	84	5	A42985	A42985 Sequence 20
36	12.2	67.8	84	5	A45530	A45530 Sequence 15
37	12.2	67.8	84	5	AR030802	AR030802 Sequence
38	12.2	67.8	94	5	I33294	I33294 Sequence 26
39	12	66.7	77	11	AF032263	AF032263 Lemur sp.
40	11.8	65.6	21	5	A60234	A60234 Sequence 1
41	11.8	65.6	30	5	A36556	A36556 Sequence 10
42	11.8	65.6	30	5	A36557	A36557 Sequence 11
43	11.8	65.6	30	5	A36558	A36558 Sequence 12
44	11.8	65.6	30	5	A60242	A60242 Sequence 9
45	11.8	65.6	38	5	A11228	A11228 Oligonucleo

ALIGNMENTS

RESULT 1
E07840/c 25 bp DNA PAT 29-SEP-1997
LOCUS Upstream activation sequence in AOX2 promoter(Alcohol Oxidase 2
DEFINITION promoter).
ACCESSION E07840
VERSION E07840.1 GI:2175973
KEYWORDS JP 1994189769-A/10.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 25)
AUTHORS Oi,H., Miura,M., Uno,S., Chuganji,M., Hiramatsu,T., and Omura,T.
TITLE VARIANT AOX2 PROMOTER, VECTOR CARRYING THE PROMOTER, TRANSFORMANT
JOURNAL AND PRODUCTION OF FOREIGN PROTEIN
PATENT: JP 1994189769-A 10 12-JUL-1994;
COMMENT GREEN CROSS CORP:THE
OS None
OC Artificial sequences.
PN JP 1994189769-A/10
PD 12-JUL-1994
PE 06-AUG-1993 JP 1993215306
PR 30-OCT-1992 JP 92P 293315
PI OI HIDEYUKI, MIURA MASAMI, UNO SHUSEI, CHUGENJI MASAKO, PI
HIRAMATSU TAKASHI,
PI OMURA TAKAO
PC C12N15/14,C12N1/19,C12N15/67,C12N15/81,C12P21/02,(C12N1/19, PC
C12R1:84)
PC (C12P21/02,C12R1:84);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..25
FT Location/Qualifiers
FEATURES
source 1..25
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 7 a 3 c 5 g 10 t
ORIGIN

Query Match 76.7%; Score 13.8; DB 5; Length 25;
Best Local Similarity 88.2%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaacagccta 17
|||||
Db 19 GCGACAAAATAGCCTA 3

RESULT 2
I72301/c 25 bp DNA PAT 03-APR-1998
LOCUS I72301
DEFINITION Sequence 10 from patent US 5683893.
ACCESSION I72301
VERSION I72301.1 GI:3008440
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 25).
AUTHORS Oh,H., Miura,M., Uno,S., Chuganji,M., Hiramatsu,R. and Ohmura,T.
TITLE Mutant AOX2 promoter, vector carrying same, transformant, and
JOURNAL Production of heterologous protein
PATENT: US 5683893-A 10 04-NOV-1997;
FEATURES Location/Qualifiers
source 1..25
/organism="unknown"
BASE COUNT 7 a 3 c 5 g 10 t
ORIGIN

Query Match 76.7%; Score 13.8; DB 5; Length 25;
Best Local Similarity 88.2%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaacagccta 17
|||||
Db 19 GCGACAAAATAGCCTA 3

RESULT 3
I79557/c 25 bp DNA PAT 10-JUN-1998
LOCUS I79557
DEFINITION Sequence 2 from patent US 5707827.
ACCESSION I79557
VERSION I79557.1 GI:3207847
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 25)
AUTHORS Oh,H., Miura,M., Hiramatsu,R. and Ohmura,T.
TITLE Mutant AOX2 promoter, vector carrying same, transformant and
JOURNAL production of heterologous protein
PATENT: US 5707827-A 2 13-JAN-1998;
FEATURES Location/Qualifiers
source 1..25
/organism="unknown"
BASE COUNT 7 a 3 c 5 g 10 t
ORIGIN

Query Match 76.7%; Score 13.8; DB 5; Length 25;
Best Local Similarity 88.2%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaacagccta 17
|||||
Db 19 GCGACAAAATAGCCTA 3

RESULT 4
I72318/c 30 bp DNA PAT 03-APR-1998
LOCUS I72318
DEFINITION Sequence 27 from patent US 5683893.
ACCESSION I72318
VERSION I72318.1 GI:3008457
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Oh,H., Miura,M., Uno,S., Chuganji,M., Hiramatsu,R. and Ohmura,T.
TITLE Mutant AOX2 promoter, vector carrying same, transformant, and
JOURNAL production of heterologous protein
PATENT: US 5683893-A 27 04-NOV-1997;
FEATURES Location/Qualifiers
source 1..30
/organism="unknown"
BASE COUNT 9 a 5 c 5 g 11 t
ORIGIN

Query Match 76.7%; Score 13.8; DB 5; Length 30;
Best Local Similarity 88.2%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaacagccta 17
|||||
Db 23 GCGACAAAATAGCCTA 7

RESULT 5
LOCUS 172311/c 31 bp DNA PAT 03-APR-1998
DEFINITION Sequence 20 from patent US 5683893.
ACCESSION 172311
VERSION 172311.1 GI:3008450
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 31)
AUTHORS Ohl,H., Miura,M., Uno,S., Chuganji,M., Hiramatsu,R. and Ohmura,T.
TITLE Mutant AOX2 promoter, vector carrying same, transformant, and production of heterologous protein
JOURNAL Patent: US 5683893-A 20 04-NOV-1997;
FEATURES Location/Qualifiers
source 1..31
/organism="unknown"
BASE COUNT 9 a 4 c 6 g 12 t
ORIGIN

Query Match 76.7%; Score 13.8; DB 5; Length 31;
Best Local Similarity 88.2%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gcgacaaacccagccta 17
|||||
Db 24 GCGACAAAATAGCCTA 8

RESULT 6
LOCUS 172312 31 bp DNA PAT 03-APR-1998
DEFINITION Sequence 21 from patent US 5683893.
ACCESSION 172312
VERSION 172312.1 GI:3008451
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 31)
AUTHORS Ohl,H., Miura,M., Uno,S., Chuganji,M., Hiramatsu,R. and Ohmura,T.
TITLE Mutant AOX2 promoter, vector carrying same, transformant, and production of heterologous protein
JOURNAL Patent: US 5683893-A 21 04-NOV-1997;
FEATURES Location/Qualifiers
source 1..31
/organism="unknown"
BASE COUNT 12 a 6 c 4 g 9 t
ORIGIN

Query Match 76.7%; Score 13.8; DB 5; Length 31;
Best Local Similarity 88.2%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gcgacaaacccagccta 17
|||||
Db 12 GCGACAAAATAGCCTA 28

RESULT 7
LOCUS 179569/c 31 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 14 from patent US 5707827.
ACCESSION 179569
VERSION 179569.1 GI:3207859
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 31)
Unclassified.

AUTHORS Ohl,H., Miura,M., Hiramatsu,R. and Ohmura,T.
TITLE Mutant AOX2 promoter, vector carrying same, transformant and production of heterologous protein
JOURNAL Patent: US 5707827-A 14 13-JAN-1998;
FEATURES Location/Qualifiers
source 1..31
/organism="unknown"
BASE COUNT 9 a 4 c 6 g 12 t
ORIGIN

Query Match 76.7%; Score 13.8; DB 5; Length 31;
Best Local Similarity 88.2%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gcgacaaacccagccta 17
|||||
Db 24 GCGACAAAATAGCCTA 8

RESULT 8
LOCUS 179570 31 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 15 from patent US 5707827.
ACCESSION 179570
VERSION 179570.1 GI:3207860
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 31)
AUTHORS Ohl,H., Miura,M., Hiramatsu,R. and Ohmura,T.
TITLE Mutant AOX2 promoter, vector carrying same, transformant and production of heterologous protein
JOURNAL Patent: US 5707827-A 15 13-JAN-1998;
FEATURES Location/Qualifiers
source 1..31
/organism="unknown"
BASE COUNT 12 a 6 c 4 g 9 t
ORIGIN

Query Match 76.7%; Score 13.8; DB 5; Length 31;
Best Local Similarity 88.2%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gcgacaaacccagccta 17
|||||
Db 12 GCGACAAAATAGCCTA 28

RESULT 9
LOCUS 172317/c 50 bp DNA PAT 03-APR-1998
DEFINITION Sequence 26 from patent US 5683893.
ACCESSION 172317
VERSION 172317.1 GI:3008456
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 50)
AUTHORS Ohl,H., Miura,M., Uno,S., Chuganji,M., Hiramatsu,R. and Ohmura,T.
TITLE Mutant AOX2 promoter, vector carrying same, transformant, and production of heterologous protein
JOURNAL Patent: US 5683893-A 26 04-NOV-1997;
FEATURES Location/Qualifiers
source 1..50
/organism="unknown"
BASE COUNT 15 a 9 c 7 g 19 t
ORIGIN

Query Match 76.7%; Score 13.8; DB 5; Length 50;
Best Local Similarity 88.2%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 gcgacaaaccagccta 17
|||||||
Db 41 GCGACAAAATAGCCTA 25

RESULT 10
A36554/c 30 bp DNA PAT 05-MAR-1997
LOCUS A36554
DEFINITION Sequence 8 from Patent W09325708.
ACCESSION A36554
VERSION A36554.1 GI:2293865
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Darrasse, A., Kocoujansky, A. and Bertheau, Y.
TITLE NUCLEOTIDIC SEQUENCES OBTAINED FROM GENES CODING PEPTATE-LYASES, AND UTILIZATIONS THEREOF PARTICULARLY FOR THE DETECTION OF BACTERIA OF THE GENUS -*(ERWINIA)*
JOURNAL Patent: WO 9325708-A 8 23-DEC-1993;
COMMENT AGRONOMIQUE INST NAT RECH (FR)
OTHER PUBLICATION CA 2137253 931223
OTHER PUBLICATION FR 2691978 931210.
FEATURES
LOCATION/Qualifiers
1..30
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 3 a 6 c 11 g 10 t
ORIGIN

Query Match 74.4%; Score 13.4; DB 5; Length 30;
Best Local Similarity 93.3%; Pred. No. 1.9e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 cgacaaccagcct 16
|||||||
Db 27 CGACAAACCACTCT 13

RESULT 11
MUSENIGO 78 bp DNA ROD 27-APR-1993
LOCUS
DEFINITION Mouse DNA upstream of Ig lambda-2 with an insertion site for transposon Etn.
ACCESSION M30932
VERSION M30932.1 GI:193188
KEYWORDS transposon.
SOURCE
ORGANISM
MUS MUSCULUS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Weiss, S. and Johansson, B.
TITLE Integration of the transposon-like element Etn upstream of V-lambda-2 in the cell line P3X63Ag8
JOURNAL J. Immunol. 143, 2384-2388 (1989)
MEDLINE 89381396
FEATURES
LOCATION/Qualifiers
1..78
/organism="Mus musculus"
/db_xref="taxon:10090"
misc.feature 43..44
BASE COUNT 33 a 13 c 15 g 17 t
ORIGIN

Query Match 73.3%; Score 13.2; DB 12; Length 78;
Best Local Similarity 83.3%; Pred. No. 2.8e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 gcgacaaaccagcctag 18
|||||||
Db 22 GCAACAAAACACCTAG 39

RESULT 12
A36555/c 30 bp DNA PAT 05-MAR-1997
LOCUS A36555
DEFINITION Sequence 9 from Patent W09325708.
ACCESSION A36555
VERSION A36555.1 GI:2293866
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Darrasse, A., Kocoujansky, A. and Bertheau, Y.
TITLE NUCLEOTIDIC SEQUENCES OBTAINED FROM GENES CODING PEPTATE-LYASES, AND UTILIZATIONS THEREOF PARTICULARLY FOR THE DETECTION OF BACTERIA OF THE GENUS -*(ERWINIA)*
JOURNAL Patent: WO 9325708-A 9 23-DEC-1993;
COMMENT AGRONOMIQUE INST NAT RECH (FR)
OTHER PUBLICATION CA 2137253 931223
OTHER PUBLICATION FR 2691978 931210.
FEATURES
LOCATION/Qualifiers
1..30
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 5 a 6 c 9 g 10 t
ORIGIN

Query Match 68.9%; Score 12.4; DB 5; Length 30;
Best Local Similarity 92.9%; Pred. No. 7.5e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 3 gacaaccagcct 16
|||||||
Db 26 GACAAACCACTCT 13

RESULT 13
A19278/c 36 bp DNA PAT 01-JUN-1994
LOCUS A19278
DEFINITION Oligonucleotide.
ACCESSION A19278
VERSION A19278.1 GI:513506
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
AUTHORS
TITLE XYLANASE PRODUCTION
JOURNAL Patent: WO 9119782-A 18 26-DEC-1991;
FEATURES
LOCATION/Qualifiers
1..36
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 6 a 8 c 6 g 16 t
ORIGIN

Query Match 67.8%; Score 12.2; DB 5; Length 36;
Best Local Similarity 82.4%; Pred. No. 1e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 gcgacaaaccagccta 17
|||||||

Db 29 GCTGCAAAACGACCA 13

RESULT 14

A36995/c 39 bp DNA PAT 05-MAR-1997

DEFINITION Sequence 46 from Patent WO9403578.

ACCESSION A36995

VERSION A36995.1 GI:2294189

KEYWORDS

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 39)

AUTHORS Hilden,H.T., Marugg,J.D., Warr,J.F., Klugkist,J., Musters,W. and Hornmann,D.H.

TITLE ENZYMATIC DETERGENT COMPOSITIONS

JOURNAL Patent: WO 9403578-A 46 17-FEB-1994;

COMMENT UNILEVER PLC (GB)

Other publication CZ 9500242 951018

Other publication CN 1088256 940622

Other publication PL 307269 950515

Other publication CA 2141559 940217

Other publication AU 4700793 940303

Other publication ZA 9305530 950130

Other publication JP 8502084T 960305.

FEATURES

source 1..39

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 8 a 7 c 11 g 13 t

ORIGIN

Query Match 67.8%; Score 12.2; DB 5; Length 39;

Best Local Similarity 82.4%; Pred. No.1e+04; 3; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gcgacaaaccagccta 17

|| |||||

Db 20 GCTGCAAAACGACCA 4

RESULT 15

A39203/c 39 bp DNA PAT 05-MAR-1997

DEFINITION Sequence 51 from Patent WO9414963.

ACCESSION A39203

VERSION A39203.1 GI:2295572

KEYWORDS

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 39)

AUTHORS Edmond,M.R., Van,D.H., Musters,W., Peters,H., Verrips,C.T. and De,V.J.

TITLE MODIFIED CUTINASES, DNA, VECTOR AND HOST

JOURNAL Patent: WO 9414963-A 51 07-JUL-1994;

COMMENT UNILEVER PLC (GB)

Other publication SK 79595 951108

Other publication PL 309388 951002

Other publication CA 2150837 940707

Other publication AU 5699994 940719

Other publication CN 1090328 940803

Other publication CZ 9501578 951213

Other publication ZA 9309415 950615.

FEATURES

source 1..39

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 8 a 7 c 11 g 13 t

ORIGIN

Query Match 67.8%; Score 12.2; DB 5; Length 39;

Best Local Similarity 82.4%; Pred. No.1e+04; 3; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gcgacaaaccagccta 17

|| |||||

Db 20 GCTGCAAAACGACCA 4

Search completed: June 4, 2000, 16:05:35

Job time: 27887 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:42 ; Search time 244.64 Seconds
(without alignments)
14.877 Million cell updates/sec

Title: US-09-164-714-11

Perfect score: 28

Sequence: 1 ctgaccaatttggcaccacacatagg 28

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/PCBUS.COMB.seq:*
7: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	57.1	39	1	US-08-361-920-32 Sequence 32, Appl
2	16	57.1	39	1	US-08-479-939-32 Sequence 32, Appl
3	16	57.1	39	2	US-08-483-433-32 Sequence 32, Appl
4	15	53.6	34	4	US-08-459-135A-2 Sequence 2, Appl
5	14.2	50.7	72	1	US-08-198-431-31 Sequence 31, Appl
6	14.2	50.7	72	2	US-08-564-955-31 Sequence 31, Appl
7	14.2	50.7	72	2	US-08-537-874-31 Sequence 31, Appl
8	14	50.0	36	2	US-08-764-100-6 Sequence 6, Appl
9	13.8	49.3	35	3	US-08-583-276-9 Sequence 9, Appl
10	13.8	49.3	35	3	US-08-583-276-10 Sequence 10, Appl
11	13.6	48.6	21	4	US-08-743-637B-210 Sequence 210, Appl
12	13.6	48.6	26	2	US-08-538-875-26 Sequence 26, Appl
13	13.4	47.9	18	5	US-09-289-376-21 Sequence 21, Appl
14	13.4	47.9	44	2	US-08-592-406-5 Sequence 5, Appl
15	13.4	47.9	97	1	US-07-635-561A-3 Sequence 3, Appl
16	13.2	47.1	23	3	US-08-359-705B-11 Sequence 11, Appl
17	13.2	47.1	23	3	US-08-286-846A-11 Sequence 11, Appl
18	13.2	47.1	23	4	US-08-457-880A-11 Sequence 11, Appl
19	13.2	47.1	23	5	US-08-444-622A-11 Sequence 11, Appl
20	13.2	47.1	23	5	US-08-942-562-11 Sequence 11, Appl
21	13.2	47.1	29	1	US-08-105-483-268 Sequence 268, Appl
22	13.2	47.1	29	2	US-08-224-391-84 Sequence 84, Appl
23	13.2	47.1	29	2	US-08-484-304-84 Sequence 84, Appl
24	13.2	47.1	29	2	US-08-224-657-105 Sequence 105, Appl
25	13.2	47.1	29	2	US-08-709-209-268 Sequence 268, Appl
26	13.2	47.1	29	2	US-08-257-073-63 Sequence 63, Appl
27	13.2	47.1	29	2	US-08-458-101-266 Sequence 268, Appl

28	13.2	47.1	29	3	US-08-184-009-134	Sequence 134, Appl
29	13.2	47.1	29	3	US-08-566-398-45	Sequence 45, Appl
30	13.2	47.1	29	4	US-08-458-356-134	Sequence 134, Appl
31	13.2	47.1	29	4	US-08-658-665-88	Sequence 88, Appl
32	13.2	47.1	43	1	US-07-885-689A-24	Sequence 24, Appl
33	13.2	47.1	43	4	US-08-857-946-116	Sequence 116, Appl
34	13.2	47.1	43	5	US-08-970-740-116	Sequence 116, Appl
35	13	46.4	21	1	US-08-434-411-38	Sequence 38, Appl
36	13	46.4	21	1	US-08-434-402-38	Sequence 38, Appl
37	13	46.4	21	2	US-08-783-288-38	Sequence 38, Appl
38	13	46.4	21	4	US-08-890-640-38	Sequence 38, Appl
39	13	46.4	21	7	5194592-69	Patent No. 5194592
40	13	46.4	30	6	PCT-US93-01598-27	Sequence 27, Appl
41	13	46.4	48	1	US-08-116-389-16	Sequence 16, Appl
42	13	46.4	48	2	US-08-708-431-16	Sequence 16, Appl
43	13	46.4	48	3	US-08-880-830-16	Sequence 16, Appl
44	13	46.4	48	6	PCT-US94-13895-16	Sequence 16, Appl
45	13	46.4	56	6	PCT-US91-05177-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
Sequence 32, Application US/08361920
Patent No. 5457046
GENERAL INFORMATION:
APPLICANT: Moeldike, Helle F.
APPLICANT: Hjoert, Carsten M.
APPLICANT: Sven, Hastrup
TITLE OF INVENTION: An Enzyme Capable of Degrading Cellulose
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 54570460 No. 5457046disk of No. 5457046th America, Inc.
STREET: 405 Lexington Avenue, 62nd Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6201
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08361, 920
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/940, 860
FILING DATE: 28-OCT-1992
APPLICATION NUMBER: DK 1158/90
FILING DATE: 09-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00124
FILING DATE: 08-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3435, 204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-867-0298
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-361-920-32

Query Match 57.1%; Score 16; DB 1; Length 39;
Best Local Similarity 79.2%; Pred. No. 47;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 4 gaccatttggcaccacatag 27
||||| ||| | ||| ||
Db 11 GACCAATTCTGGCCACACATTGG 34

RESULT 2
US-08-479-939-32

Sequence 32, Application US/08479939
Patent No. 5686593

GENERAL INFORMATION:

APPLICANT: Woeldike, Helle F.

APPLICANT: Hagen, Frederick

APPLICANT: Hjort, Carsten M.

APPLICANT: Sven, Hasstrup

TITLE OF INVENTION: An Enzyme Capable of Degrading Cellulose

TITLE OF INVENTION: or Hemicellulose

NUMBER OF SEQUENCES: 85

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5686593o No. 5686593disk of No. 5686593th America, Inc.

STREET: 405 Lexington Avenue, 62nd Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/479, 939

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/361, 920

FILING DATE: 22-DEC-1994

APPLICATION NUMBER: US 07/940, 860

FILING DATE: 28-OCT-1992

APPLICATION NUMBER: DK 1158/90

FILING DATE: 09-MAY-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/DK91/00124

FILING DATE: 08-MAY-1991

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 3435, 204-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-867-0298

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 39 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-479-939-32

Query Match 57.1%; Score 16; DB 1; Length 39;
Best Local Similarity 79.2%; Pred. No. 47;

Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 4 gaccatttggcaccacatag 27
||||| ||| | ||| ||
Db 11 GACCAATTCTGGCCACACATTGG 34

RESULT 3
US-08-483-432-32

Sequence 32, Application US/08483432
Patent No. 5763254

GENERAL INFORMATION:

APPLICANT: Woeldike, Helle F.

APPLICANT: Hagen, Frederick

APPLICANT: Hjort, Carsten M.

APPLICANT: Sven, Hasstrup

TITLE OF INVENTION: An Enzyme Capable of Degrading Cellulose

TITLE OF INVENTION: or Hemicellulose

NUMBER OF SEQUENCES: 85

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5763254o No. 5763254disk of No. 5763254th America, Inc.

STREET: 405 Lexington Avenue, 62nd Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/483, 432

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/361, 920

FILING DATE: 28-OCT-1992

APPLICATION NUMBER: DK 1158/90

FILING DATE: 09-MAY-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/DK91/00124

FILING DATE: 08-MAY-1991

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 3435, 204-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-867-0298

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 39 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-483-432-32

Query Match 57.1%; Score 16; DB 2; Length 39;
Best Local Similarity 79.2%; Pred. No. 47;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 4 gaccatttggcaccacatag 27
||||| ||| | ||| ||
Db 11 GACCAATTCTGGCCACACATTGG 34


```

:
: REGISTRATION NUMBER: 34,587
: REFERENCE/DOCKET NUMBER: 1652857-014611US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2400
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 31:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 72 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-564-955-31

```

[illegible]

```

RESULT 7
US-08-537-874-31
Sequence 31, Application US/08537874
Patent No. 5830721
GENERAL INFORMATION:
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Cramerli, Andreas
TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation
TITLE OF INVENTION: and Reassembly
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/537,874
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP PCT/US95/02126
FILING DATE: 17-FEB-1995
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018097-014610
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ. ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 72 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-537-874-31
Query Match 50.7%; Score 14.2; DB 3; Length 72;
Best Local Similarity 70.4%;
Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 8; Indels

```

Qy	1	cttgaccaatttggaccaacatag	27
Db	39	CGTTACATATATTGGTACCAACAGAAG	65

RESULT 8
US-08-764-100-6/c
; sequence 6, Application US/08764100
; Patent No. 5773700

```

APPLICANT: van Grinsven J., Martinus Q.
APPLICANT: de Haan, Petrus T.
APPLICANT: Gielen L., Johannes J.
APPLICANT: Peters, Dirk
APPLICANT: Goldbach, Robert W.
TITLE OF INVENTION: Improvements in or Relating to Organoid
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,100
FILING DATE: 06-DEC-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,064
FILING DATE:
APPLICATION NUMBER: US 08/032,235
FILING DATE: 17-MAR-1993
APPLICATION NUMBER: GB 9206016.9
FILING DATE: 19-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5737001s, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER: 137-1061
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 354-3592
TELEFAX: (415) 857-1125
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IS-08-764-100-6

```

```

Query Match      50.0%  Score 14; DB 2; Length 36;
Best Local Similarity 77.3%  Pred. NO. 3.7e+02;
Matches 17; Conservative 0; Mismatches 5; Indels 0;
                                Gaps 0;

OY      4  gaccatttggcaccacata 25
          || ||||| || || || ||
db      35  GAGCAATTGTGTCCAGGATATA 14

```

RESULT 9
US-08-583-276-9/c
; Sequence 9, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonald, Kevin T.

```

;
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DM4.V2
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/332,444
; FILING DATE: 31-OCT-1994
; APPLICATION NUMBER: 07/887,712
; FILING DATE: 22-MAY-1992
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 bases
; TYPE: nucleic acid
; STRANDEDNESS: singular
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: DNA primer
;
; US-08-583-276-9
;
;
; Query Match          49.3%; Score 13.8; DB 3; Length 35;
; Best Local Similarity 88.2%; Pred. No. 4.5e+02;
; Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 7 caatttgcaccacaca 23
;      ||||||| ||||| |
; DB 35 CAATTGTGCACCAATA 19
;
; RESULT 10
; US-08-583-276-10
; Sequence 10, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonald, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DM4.V2

```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/332,444
; FILING DATE: 31-OCT-1994
; APPLICATION NUMBER: 07/887,712
; FILING DATE: 22-MAY-1992
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 bases
; TYPE: nucleic acid
; STRANDEDNESS: singular
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: DNA primer
;
; US-08-583-276-10
;
;
; Query Match          49.3%; Score 13.8; DB 3; Length 35;
; Best Local Similarity 88.2%; Pred. No. 4.5e+02;
; Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 7 caatttgcaccacaca 23
;      ||||||| ||||| |
; DB 1 CAATTGTGCACCAATA 17
;
; RESULT 11
; US-08-743-637B-210
; Sequence 210, Application US/08743637B
; Patent No. 5994066
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: PICARD, Francois J.
; APPLICANT: OUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: O'CARLES & BRADY
; STREET: 411 EAST WISCONSIN AVENUE
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,637B
; FILING DATE: 04-NOV-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,840
; FILING DATE: 11-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 850586, 90012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 210:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid

```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-743-637B-210

Query Match 48.6%; Score 13.6; DB 4; Length 21;
Best Local Similarity 80.0%; Pred. No. 5.1e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 tgaccaatttggcaccac 22
||||| ||||| || |||||
DB 1 TGACCACCTTTATCAGCAAC 20

RESULT 12
US-08-538-875-26/c
Sequence 26, Application US/08538875
Patent No. 5773582
GENERAL INFORMATION:
APPLICANT: Shin, Hang-Cheol
APPLICANT: Shin, Nam-Kyu
APPLICANT: Lee, Inkyung
APPLICANT: Kang, Sungzong
TITLE OF INVENTION: TUMOR NECROSIS FACTOR MUTAINS
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shin, Hang-Cheol
STREET: Jukong Gochung Apt. 1014-806, Haan-dong
CITY: Kwangmyung-shi
STATE: Kyungki-do
COUNTRY: Republic of Korea
ZIP: 423-060
ADDRESSEE: Shin, Nam-Kyu
STREET: #181-404 Sadang-4-dong, Dongjak-ku
CITY: Seoul
STATE:
COUNTRY: Republic of Korea
ZIP: 156-094
ADDRESSEE: Lee, Inkyung
STREET: 11/2, #302-39 Juan-4-dong, Nam-ku
CITY: Incheon
STATE:
COUNTRY: Republic of Korea
ZIP: 402-204
ADDRESSEE: Kang, Sungzong
STREET: #84-4 Daeshin-dong, Seodaemun-ku
CITY: Seoul
STATE:
COUNTRY: Republic of Korea
ZIP: 120-160
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5inch 2.0mb storage
COMPUTER: IBM PC/AT
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/538, 875
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/193, 336
FILING DATE:
APPLICATION NUMBER: KR 93-1751
FILING DATE: 9-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE:
TELEFAX:
INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:
LENGTH: 26 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: primer DNA
US-08-538-875-26

Query Match 48.6%; Score 13.6; DB 2; Length 26;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 4 gaccaatttggcaccacaa 23
||| ||||| ||||| ||
DB 20 GCCCTTTTGGCACCACCA 1

RESULT 13
US-09-289-376-21
Sequence 21, Application US/09289376
Patent No. 6013788
GENERAL INFORMATION:
APPLICANT: Brett P. Monia
APPLICANT: Lex M. Cowser
TITLE OF INVENTION: ANTISENSE MODULATION OF SMAD3 EXPRESSION
FILE REFERENCE: RTS-0043
CURRENT APPLICATION NUMBER: US/09/289,376
CURRENT FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 47
SEQ ID NO 21
LENGTH: 18
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
US-09-289-376-21

Query Match 47.9%; Score 13.4; DB 5; Length 18;
Best Local Similarity 93.3%; Pred. No. 6e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 13 tggcaccacacatag 27
||||| ||||| |||||
DB 4 tggcaccacacacag 18

RESULT 14
US-08-592-406-5/c
Sequence 5, Application US/08592406
Patent No. 5821059
GENERAL INFORMATION:
APPLICANT: MINION, F. Chris
APPLICANT: KNUDSON, Kevin L.
TITLE OF INVENTION: MYCOPLASMA EXPRESSION SYSTEM
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,406
FILING DATE: 06-FEB-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US93/07407
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 76645/132
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-592-406-5

Query Match 47.9%; Score 13.4; DB 2; Length 44;
Best Local Similarity 93.3%; Pred. No. 7.2e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ctggaccaatttgg 15
|||||||||
DB 33 CTGGACCAATTGGG 19

RESULT 15
US-07-635-561A-3/C
Sequence 3, Application US/07635561A
Patent No. 5244805
GENERAL INFORMATION:
APPLICANT: Miller, Lois K.
TITLE OF INVENTION: Improved Baculovirus Expression Vectors
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/635,561A
FILING DATE: 19910117
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/353,847
FILING DATE: 17-MAY-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 90/02814
FILING DATE: 17-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 4-90A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
TELEX: 823189
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
US-07-635-561A-3

Query Match 47.9%; Score 13.4; DB 1; Length 97;
Best Local Similarity 73.9%; Pred. No. 8.4e+02;
Matches 17; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 tgaccaatttggcaccacata 25
| ||||| ||||| ||| |||
DB 25 TCACCAAGCTGTGGCTCCATGATA 3

Search completed: June 4, 2000, 16:09:42
Job time: 28061 sec

4 . . 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:12 ; Search time 4521.53 Seconds
(Without alignments)
25.100 Million cell updates/sec

Title: US-09-164-714-11

Perfect score: 28
Sequence: 1 cttgaccaatttgccaccaacataagg 28

Scoring table:

IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 485736 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

COMMENT	FEATURES	SOURCE
<p>On May 9, 1995 this sequence version replaced gi:803068.</p> <p>Contact: Marra M/Mouse EST Project</p> <p>WashU-HHMI Mouse EST Project</p> <p>Washington University School of Medicine</p> <p>4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108</p> <p>Tel: 314 286 1800</p> <p>Fax: 314 286 1810</p> <p>Email: mouseest@watson.wustl.edu</p> <p>This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.</p> <p>MC1:490258</p>	<p>Trace considered overall poor quality</p> <p>Possible reversed clone: similarity on wrong strand</p> <p>Seq primer: -28ml3 rev1 ET from Amersham</p> <p>High quality sequence stop: 1.</p> <p>location/Qualifiers</p> <p>1. 67</p> <p>/organism="Mus musculus"</p> <p>/strain="C57BL6 x DBA"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="IMAGE:821978"</p> <p>/clone_1lb="Beddington mouse embryonic region"</p> <p>/sex="pooled"</p> <p>/tissue-type="embryo"</p> <p>/dev_stage="7.5dpc"</p> <p>/lab_host="DH12S"</p> <p>/note="Organ: whole embryo; Vector: pCMV-SPORT; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt. Gastrulating embryos were collected at 7.5dpc from C57BL6 x DBA matings, excluding embryos that had developed head folds and all extraembryonic tissues. Average insert size: 1.3 kb (range: 0.5 - 3.0 kb). Referenced in Development 121, 2479-2489 (1995)."</p>	<p>BASE COUNT</p> <p>19 a 23 c 13 g 12 t</p> <p>ORIGIN</p>
<p>Query Match</p> <p>Best Local Similarity 81.8%; Pred. No. 3,1e+03;</p> <p>Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;</p>	<p>55.7%; Score 15.6; DB 34; Length 67;</p>	<p>QY</p> <p>5 accaatlttgacacacaatag 26</p> <p> </p> <p>Db 25 ACTAAATTGGCTCAGCATAG 4</p>
RESULT	2	

ACCESSION	AI209047	
VERSION	AI209047.1	GI:3770989
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 46)	
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .	
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	
COMMENT	Tumor Gene Index	
	Unpublished (1997)	
	On Jan 19, 1998 this sequence version replaced gi:2151684.	
	Contact: Robert Strausberg, Ph.D.	
	Tel: (301) 496-1550	
	Email: Robert_Strausberg@nih.gov	
	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.	
	Emmert-Buck, M.D., Ph.D.	
	cDNA Library Preparation: M. Bento Soares, Ph.D.	
	cDNA Library Arrayed by: Greg Lennon, Ph.D.	
	DNA Sequencing by: Washington University Genome Sequencing Center	

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: www.bio.lnl.gov/db/rp/image/image.html

Trace considered overall poor quality
Insert Length: 692 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers

FEATURES

source

1. .46
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:176239A"
/clone_lib="NCI_CGAP_Kid3"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer, double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. mRNA source: 2 pooled kidneys. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 10 a 9 c 17 g 10 t
ORIGIN

Query Match 54.3%; Score 15.2; DB 43; Length 46;
Best Local Similarity 85.0%; Pred. No. 4.4e+03;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 cttgaccaatttggcaccac 20
||| ||| ||| ||| ||| |||
Db 37 CTTTACCACTTTTGCACCA 18

RESULT 3
A1739189 73 bp mRNA EST 18-JUN-1999
LOCUS A1739189
DEFINITION w127g11.x1 NCI_CGAP Col6 Homo sapiens cDNA clone IMAGE:2391476 3'
similar to gb:U1206 40S RIBOSOMAL PROTEIN S4 (HUMAN);, mRNA
sequence.

ACCESSION A1739189
VERSION A1739189.1 GI:5101170
KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 73)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

Unpublished (1997)

On Jun 5, 1998 this sequence version replaced gi:3189015.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
www.bio.lnl.gov/db/rp/image/image.html

Seq primer: -40UP from Gibco.

FEATURES

source

1. .73
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:2391476"
/clone_lib="NCI_CGAP_Col6"
/tissue_type="colon tumor, RER+"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; plasmid DNA from the normalized library NCI_CGAP_Col10 was prepared, and ss circles were made in vitro. Following BAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1057416-1061255, and 1144584-1145351).
Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 17 a 18 c 13 g 25 t
ORIGIN

Query Match 53.6%; Score 15; DB 51; Length 73;
Best Local Similarity 78.3%; Pred. No. 5.9e+03;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 6 ccaattggcaccacatagg 28
||| ||| ||| ||| ||| |||
Db 58 CCAGCTGTGCTACACATAGG 36

RESULT 4
A1745436 82 bp mRNA EST 21-JUN-1999
LOCUS A1745436
DEFINITION tr04g11.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2217380 3'
similar to gb:U1206 40S RIBOSOMAL PROTEIN S4 (HUMAN);, mRNA
sequence.

ACCESSION A1745436
VERSION A1745436
KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 82)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

Unpublished (1997)

On May 18, 1998 this sequence version replaced gi:3137118.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
www.bio.lnl.gov/db/rp/image/image.html

Seq primer: -40UP from Gibco.

FEATURES

source

1. .82
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2217380"
/clone_lib="NCI_CGAP_Ov23"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog #: 11534-013"

BASE COUNT 17 a 15 c 10 g 40 t


```

VERSION      AA714476.1  GI:2726750
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS      Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        1 (bases 1 to 97)
JOURNAL      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
              Unpublished (1997)
              On Sep 12, 1996 this sequence version replaced gi:1407494.
              Contact: Robert Strausberg, Ph.D.
              Tel: (301) 496-1550
              Email: Robert_Strausberg@nih.gov
              Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
              Emmerit-Buck, M.D., Ph.D.
              CDNA Library Preparation: Stratagene, Inc.
              CDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LINL at:
              www.bio.lnlnl.gov/db/brp/image/image.html

BASE COUNT   18 a      24 c      26 g      29 t
ORIGIN
Query Match          53.6%; Score 15; DB 37; Length 97;
Best Local Similarity 78.3%; Pred. No. 6.2e+03;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 ccaatttgacacacatagg 28
    ||| 1 |||| |||||
Db 47 CCAGCTGTGCTACACATAGG 25

RESULT 8
AA566933/c 59 bp mRNA EST 28-AUG-1998
LOCUS      987 lobiolly pine CA Pinus taeda cDNA clone 1CA12G, mRNA sequence.
DEFINITION AA566933
ACCESSION  AA566933
VERSION     AA566933.1 GI:3366146
KEYWORDS    EST.
SOURCE      1obiolly pine.
ORGANISM    Pinus taeda
REFERENCE    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
              Pinaceae; Pinus.
              1 (bases 1 to 59)
AUTHORS      Altona,T., Quinn,M., Shoop,E., Swope,K., St.Cyr,S., Carlis,J.,
              Riedl,J., Retzel,E., Campbell,M.W., Sederoff,R. and Whetten,R.W.
              Analysis of xylem formation in pine by cDNA sequencing
              Proc. Natl. Acad. Sci. U.S.A. 95 (16), 9693-9698 (1998)
JOURNAL      98356220
MEDLINE
On Sep 12, 1996 this sequence version replaced gi:1394808.
COMMENT      Contact: Ross Whetten

```

```

Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall,
Raleigh, NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhet@unity.ncsu.edu
Seq primer: 73.

FEATURES
source
Location/Qualifiers
1..59
/organism="Pinus taeda"
/strain="Coastal plain lobiolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="1CA12G"
/clone_lib="lobiolly pine CA"
/tissue_type="xylem"
/lab_host="SOLR"
/note="Vector: lambda-ZAP; Site.1: EcoRI; Site.2: XhoI;
The result of subtraction of C library with N library.
Immature xylem from the underside of inclined stems of
differentiating compression wood was subtracted with
immature xylem from the side of inclined stems of
differentiating wood. A mixture of four genotypes were
used. Oligo-dT primed cDNA was directionally cloned into
the EcoRI-XhoI lambda-ZAP vector arms"

BASE COUNT   15 a      12 c      16 g      14 t      2 others
ORIGIN
Query Match          52.9%; Score 14.8; DB 35; Length 59;
Best Local Similarity 84.2%; Pred. No. 7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ttgacatttgacacca 20
    ||| 1 ||| |||||
Db 32 TTGATCATTTTGGCACCA 14

RESULT 9
AI795116/c 87 bp mRNA EST 13-DEC-1999
LOCUS      5B76H05.Y1 Gm-cl010 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl010-946 5', mRNA sequence.
ACCESSION  AI795116
VERSION     AI795116.1 GI:5342832
KEYWORDS    EST.
SOURCE      soybean.
ORGANISM    Glycine max
REFERENCE    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
              eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
              Glycine.
              1 (bases 1 to 87)
AUTHORS      Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Coryell,V.,
              Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
              Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
              Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
              Schuck,R., Ritter,E., Kohn,S., Shiu,T., Jackson,Y., Cardenas,M.,
              McCann,R., Waterston,R. and Wilson,R.
              Public Soybean EST Project
              Unpublished (1999)
              On Jun 22, 1998 this sequence version replaced gi:3247195.
              Contact: Shoemaker R/Public Soybean EST Project
              Public Soybean EST Project
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              This clone is available through: Genome Systems, Inc. 4633 World
              Parkway Circle St. Louis, Missouri 63134 For further information
              call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
              427-3324 or contact: clones@genomesystems.com or

```


/lab_host="DH10B"
/note="vector: pbluescript SK-; Site_1: EcoRI; Site_2: XhoI; Construction of the cDNA library was carried out using Stratagene 'Unizap - cDNA synthesis kit'. cDNA was constructed using an oligo dt primer/linker that contains a XhoI site within it. Following ds cDNA synthesis, EcoRI adapters were ligated to the blunt ends and sample was digested with XhoI. The result is cDNA with an EcoRI sticky end on one side and a XhoI sticky end on the other. This cDNA was ligated directionally in Unizap arms. The vector is designed containing the pbluescript sequence as well as lambda DNA and cDNA is cloned within this pbluescript sequence. The vector was then packaged using Gold gigaPackaging extracts. Library was grown in XLBlue MRF cells and amplified. The library was excised by mass excision using Stratagene 'Mass excision kit' that uses exsist as a helper phage that releases the pbluescript sequence and circularises it as single stranded plasmids that are then packaged (by helper phage) and secreted out of the host cell as phagemids. SOLR cells were transformed with phagemids and the library was plated out on LB-amp plates to select for transformants. Approximately 1,000,000 colonies were grown and recovered. The double stranded plasmid library was recovered by using Qulagen MidI prep kit. 2 micro grams of each library were used to transform DH10B cells by electroporation."

BASE COUNT
ORIGIN
27 a 23 c 24 g 25 t

Query Match 52.1%; Score 14.6; DB 70; Length 99;
Best Local Similarity 81.0%; Pred. No. 9.4e+03;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 cttgaccaatttgcaccaa 21
||||| ||||| ||||| |||||
Db 69 cttggtcattctgtcctcaa 89

RESULT 14
AA165762 79 bp mRNA EST 12-FEB-1997
LOCUS ms06f11.r1 StrataGene mouse embryonic carcinoma (#937317) Mus
DEFINITION musculus cDNA clone IMAGE:615981 5' similar to TR:E93245 E93245 ETN
INSERT IN THE FAS APOPTOSIS GENE OF MRL-IPR/IPR. [1] ; mRNA
SEQUENCE.
ACCESSION AA165762
VERSION AA165762.1 GI:1743977
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muride; Murinae; Mus.
REFERENCE 1 (bases 1 to 79)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HM Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1325676.
Contact: Marra M/Mouse EST Project
WashU-HM Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG:376805
Seq primer: -28m13 rev1 ET from Amersham

FEATURES
SOURCE
High quality sequence stop: 75.
Location/Qualifiers
1..79
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="615981"
/clone_id="StrataGene mouse embryonic carcinoma
(#937317)"
/tissue_type="carcinoma"
/dev_stage="embryonic"
/lab_host="SOLR (kanamycin resistant)"
/note="vector: pbluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. P19 cell
line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
adaptor sequence: 5' GAATTCGGCAGG 3' ~3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTT 3' "

BASE COUNT
ORIGIN
20 a 19 c 26 g 14 t

Query Match 51.4%; Score 14.4; DB 29; Length 79;
Best Local Similarity 75.0%; Pred. No. 1.1e+04;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 accaatattggcaccacacatagag 28
||||| ||||| ||||| |||||
Db 33 AACCAATTTTGGCGCCAGAACTCGG 56

RESULT 15
AW394554 91 bp mRNA EST 07-FEB-2000
LOCUS sh32e03.v1 Gm-c1017 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-c1017-3413 5', mRNA sequence.
ACCESSION AW394554
VERSION AW394554.1 GI:6912966
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
Glycine.
REFERENCE 1 (bases 1 to 91)
AUTHORS Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Pearson,B., Swaller,T., Gibbons,M., Page,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT On Jul 9, 1999 this sequence version replaced gi:5433560.
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Seq primer: -40RP from Gibco.
Location/Qualifiers
1..91
/organism="Glycine max"
/db_xref="taxon:3847"
/clone_image="GENOME SYSTEMS CLONE ID: Gm-c1017-3413"
/clone_id="Gm-c1017"
/tissue_type="vegetable buds of field grown plants"

FEATURES
SOURCE


```
/lab host="XL10-Gold"
/note="Vector: pBluescript II XR, Site_1: EcoRI, Site_2:
XhoI. This cDNA library was constructed from mRNA isolated
from vegetable buds of field grown plants. The cDNA
library was prepared using the Stratagene pBluescript II
XR library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pBluescript vector. The ligated cDNA fragments were
transformed into XL10-Gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John
Erpeiding."
```

```
BASE COUNT      22 a      33 g      25 t
ORIGIN          11 c
```

```
Query Match      51.4%; Score 14.4; DB 81; Length 91;
Best Local Similarity 75.0%; Pred. No. 1.1e+04;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Oy      5 accaatttgcacacaacataggg 28
        | | | | | | | | | | | | | |
Db      67 AACCATTTATTCACCAACAAGGG 44
```

Search completed: June 4, 2000, 13:53:12
Job time: 20936 sec

THIS PAGE BLANK (USPTO)

AC V43049;

DT 23-OCT-1998 (first entry)
DE 5' PCR primer Deg-A used to amplify nitric oxide synthase DNA.
KW Nitric oxide synthase; treatment; epidermal; dermal condition;
KM unwanted pigmentation; PCR primer; ss.
OS Synthetic.
OS Mus sp.
OS Rattus sp.
OS Homo sapiens.
PN WO983379-A1.
PD 06-AUG-1998.
PF 03-FEB-1998; U01891.
PR 04-FEB-1997; US-037098.
PA (GENO) GEN HOSPITAL CORP.
PI Lerner EA, Lerner LH, Qureshi AA;
DR WPI: 98-437056/37.
PT Treatment of unwanted epidermal or dermal conditions - comprising
PR administration of treatment which modulates level of nitric oxide in
PT skin
PS Disclosure: Page 17, 35pp; English.
CC PCR primers V43049-50 were used to amplify DNA encoding nitric oxide
CC synthase (inducible isoform, the isoform found in nerve cells, and
CC isoform found in epithelial cells). The specification describes
CC a method for treating a subject for an unwanted epidermal or dermal
CC condition, which comprises administration of a compound which modulates
CC the level of nitric oxide in the skin. Where the condition is
CC characterised by unwanted cells or unwanted pigmentation, the method
CC includes increasing the level of nitric oxide in the skin. Where the
CC condition is characterised by the lack of or by an insufficient number
CC of dermal or epidermal cells or a lack of pigmentation, the method
CC includes reducing the level of nitric oxide in the skin.
SQ Sequence 29 Bp; 12 A; 7 C; 5 G; 2 T;

Query Match 54.3%; Score 15.2; DB 1; Length 29;
Best Local Similarity 77.3%; Pred. No. 2.4e+02;
Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 7 ccaatttggcaccacaataggg 28
||| | : ||||| |||||
DB 6 CAAGTAYGCCACCAACAAAGG 27

RESULT 3
OY1835/c
ID Q71835 standard; DNA: 34 Bp.
AC Q71835;
DT 30-MAR-1995 (first entry)
DE Primer F1B2.
KM Fibronectin binding protein; Fbp; fibronectin binding domain;
KM monoclonal antibody; Mab; adhesion; Gram-positive bacterium;
KM bacteremia; Escherichia coli; polymerase chain reaction; PCR;
KM amplification; primer; Staphylococcus aureus; ss.
OS Synthetic.
PN WO9418327-A.
PD 18-AUG-1994.
PF 04-FEB-1994; G00215.
PR 05-FEB-1993; GB-002289.
PR 20-OCT-1993; GB-021592.
PA (SMK) SMITHKLINE BEECHAM PLC.
PI Burnham MKR, Chopra I, Critchley IA, Knowles DDC;
DR WPI: 94-279748/34.
PT Fibronectin binding protein and monoclonal antibodies specific
PT for Fbp - useful to prevent adherence of Gram-positive bacteria
PT to indwelling devices or wounds
PS Disclosure: Page 29; 40pp; English.
CC Polypeptides corresponding to residues G709-T886 plus P1VPPPT,
CC G709-P838(P838T) and G709-P838 (R880S-07, respectively) of
CC S. aureus J2385 (NCIMB 40532) Fbp type A fibronectin binding
CC domain D1-D4 region (R5808) were expressed in E. coli BL21(DE4)
CC and used to raise Mabs specific for Fbp. DNA encoding
CC D1-D4(G709-T886) was obtained by PCR amplification of S. aureus
CC chromosomal DNA using primers F1B1 (Q71834) and F1B2 (Q71835).
SQ Sequence 34 Bp; 5 A; 6 C; 14 G; 9 T;

Query Match 53.6%; Score 15; DB 1; Length 34;
Best Local Similarity 78.3%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 6 ccaatttggcaccacaataggg 28
|||| | ||||| |||||
DB 29 CCAATGTCGCACCAACGTAAAG 7

RESULT 4
T18304/c
ID T18304 standard; DNA: 34 Bp.
AC T18304;
DT 13-NOV-1996 (first entry)
DE Primer for amplifying fibronectin binding domains.
KM Fibronectin; oral hygiene; gingivitis; gum disease; adhesion;
KM periodontitis; oral pathogen; caries; calculus; candidiasis;
KM oral surgery; tissue regeneration; irrigation; toothpaste;
KM dentifrice; mouthwash; lozenge; ss.
OS Synthetic.
PN WO9604003-A1.
PD 15-FEB-1996.
PF 18-JUL-1995; E02825.
PR 05-AUG-1994; GB-015902.
PA (SMK) SMITHKLINE BEECHAM PLC.
PI Barnett P, Critchley IA, Dodd I;
DR WPI: 96-129122/13.
PT Prevention of adherence of oral pathogens in the oral cavity,
PT partic. tooth surfaces - by application of a fibronectin binding
PT protein or polypeptide or a monoclonal antibody or fragment against
PT it
PS Example 1: Page 18; 41pp; English.
CC Adherence of oral pathogens, particularly to tooth surfaces, can be
CC prevented by application of a fibronectin binding protein or
CC polypeptide. The fibronectin binding protein or polypeptide is
CC useful in the manufacture of oral hygiene compositions, eg.
CC toothpaste, liquid dentifrice, mouthwash or lozenge. They are
CC useful to prevent the plaque-related development of carious lesions,
CC gingivitis, calculus or periodontal disease and to combat oral cavity
CC infections, e.g. candidiasis. They may also be useful in oral
CC surgery e.g. in guided tissue regeneration procedures to prevent
CC subsequent bacterial infection, and for irrigation of periodontal
CC pockets. Two primers (T18303, T18304) were used to amplify DNA
CC encoding the fibronectin binding domains of fibronectin binding
CC protein from the chromosomal DNA of Staphylococcus aureus J2385.
SQ Sequence 34 Bp; 5 A; 6 C; 14 G; 9 T;

Query Match 53.6%; Score 15; DB 1; Length 34;
Best Local Similarity 78.3%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 6 ccaatttggcaccacaataggg 28
|||| | ||||| |||||
DB 29 CCAATGTCGCACCAACGTAAAG 7

RESULT 5
X27477/c
ID X27477 standard; DNA: 54 Bp.
AC X27477;
DT 08-JUN-1999 (first entry)
DE Oligo 1131 for hcg alpha-subunit analogues.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation; ss.
OS Synthetic.
OS Homo sapiens.
PN WO9658957-A2.
PD 30-DEC-1998.

PF 25-JUN-1998; U13070.
PR (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI: 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
or hTSH, have an intersubunit disulphide crosslink between the
alpha- and beta-subunits to improve stability
PS Example 1; Page 60; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
(hCG), human luteinising hormone (hLH), human follicle stimulating
hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
mutens, which are modified to contain an intersubunit cysteine bond,
between an alpha-subunit cysteine and a beta-subunit cysteine, for
improved stability, the analogue retaining at least a portion of the
bioactivity for the corresponding native GPH receptor. Primers
CC X27450-X27484 were used to introduce the mutations in the hCG sequence.
CC The improved analogues are designed specifically to reduce perturbation
of the 3-dimensional structure of the hormone, thereby creating greater
likelihood that the dimer will be formed in vivo and the formed dimer
will have affinity for the native receptors and have agonistic activity
on them. The changes stabilise the GPHs and prolong the biological
activities of the hormones. The analogues can have uses as for the
native GPHs.
SQ Sequence 34 BP; 18 A; 11 C; 12 G; 13 T;
Query Match 51.4%; Score 14.4; DB 1; Length 54;
Best Local Similarity 75.0%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 tgaccatttggcaccacacatag 26
||||| ||||| ||||| ||||| |||||
DB 43 TGACGTCCTTTGGACCACACATAG 20
RESULT 6
X27467/C
ID X27467 standard; DNA; 60 BP.
AC X27467;
DR 08-JUN-1999 (first entry)
DE Oligo 850 for hCG alpha-subunit analogues.
KW Analogue: heterodimeric: glycoprotein hormone; hCG; hLH; hFSH; hTSH;
human chorionic gonadotropin; human luteinising hormone; disulphide bond;
human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability: primer; amplification; PCR; mutation; ss.
OS Synthetic.
OS Homo sapiens.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR 25-JUN-1997; US-050784.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI: 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
or hTSH, have an intersubunit disulphide crosslink between the
alpha- and beta-subunits to improve stability
PS Example 1; Page 60; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
(hCG), human luteinising hormone (hLH), human follicle stimulating
hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
mutens, which are modified to contain an intersubunit disulphide bond,
between an alpha-subunit cysteine and a beta-subunit cysteine, for
improved stability, the analogue retaining at least a portion of the
bioactivity for the corresponding native GPH receptor. Primers
CC X27450-X27484 were used to introduce the mutations in the hCG sequence.
CC The improved analogues are designed specifically to reduce perturbation
of the 3-dimensional structure of the hormone, thereby creating greater

CC likelihood that the dimer will be formed in vivo and the formed dimer
will have affinity for the native receptors and have agonistic activity
on them. The changes stabilise the GPHs and prolong the biological
activities of the hormones. The analogues can have uses as for the
native GPHs.
SQ Sequence 60 BP; 18 A; 13 C; 14 G; 15 T;
Query Match 51.4%; Score 14.4; DB 1; Length 60;
Best Local Similarity 75.0%; Pred. No. 6e+02;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 tgaccatttggcaccacacatag 26
||||| ||||| ||||| ||||| |||||
DB 43 TGACGTCCTTTGGACCACACATAG 20
RESULT 7
X27464/C
ID X27464 standard; DNA; 84 BP.
AC X27464;
DR 08-JUN-1999 (first entry)
DE Oligo 839 for hCG alpha-subunit analogues.
KW Analogue: heterodimeric: glycoprotein hormone; hCG; hLH; hFSH; hTSH;
human chorionic gonadotropin; human luteinising hormone; disulphide bond;
human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability: primer; amplification; PCR; mutation; ss.
OS Synthetic.
OS Homo sapiens.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR 25-JUN-1997; US-050784.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI: 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
or hTSH, have an intersubunit disulphide crosslink between the
alpha- and beta-subunits to improve stability
PS Example 1; Page 60; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
(hCG), human luteinising hormone (hLH), human follicle stimulating
hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
mutens, which are modified to contain an intersubunit disulphide bond,
between an alpha-subunit cysteine and a beta-subunit cysteine, for
improved stability, the analogue retaining at least a portion of the
bioactivity for the corresponding native GPH receptor. Primers
CC X27450-X27484 were used to introduce the mutations in the hCG sequence.
CC The improved analogues are designed specifically to reduce perturbation
of the 3-dimensional structure of the hormone, thereby creating greater
likelihood that the dimer will be formed in vivo and the formed dimer
will have affinity for the native receptors and have agonistic activity
on them. The changes stabilise the GPHs and prolong the biological
activities of the hormones. The analogues can have uses as for the
native GPHs.
SQ Sequence 84 BP; 25 A; 20 C; 16 G; 23 T;
Query Match 51.4%; Score 14.4; DB 1; Length 84;
Best Local Similarity 75.0%; Pred. No. 6.3e+02;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 tgaccatttggcaccacacatag 26
||||| ||||| ||||| ||||| |||||
DB 67 TGACTTCCTTTGGACCACACATAG 44
RESULT 8
097403
ID 097403 standard; DNA; 72 BP.
AC 097403;

CC 185 mutation of the mdrl gene and will differ from the sequence of
CC GIMDI only at the point mutation wherein codon 185 is changed from
CC GGT to GGT. This difference corrects the point mutation in the cDNA
CC sequence and encodes Gln in the final construction. The third set of
CC primers flank the BstXI and MscI restriction sites. These primers
CC 5' Bst and 3' Msc (Q52719, Q52720) overlap the BstXI and MscI sites,
CC respectively, and also incorporate another flanking restriction site
CC that will allow the PCR fragment to be conveniently subcloned.
CC Using the technique of recombinant PCR three separate PCR reactions
CC will be initiated, using MD1NA as the template DNA.
SQ Sequence 35 BP; 11 A; 8 C; 3 G; 13 T;

Query Match 49.3%; Score 13.8; DB 1; Length 35;
Best Local Similarity 88.2%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 caatttggcaccaca 23
DB 1 CAATTTGTCACCAATA 17

RESULT 14
T24754
ID T24754 standard; cDNA to mRNA; 67 BP.
AC T24754;
DT 09-OCT-1996 (first entry)
DE Human gene signature HUMGS06829.
KM Gene signature: messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1687; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridize to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridize with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 67 BP; 26 A; 13 C; 7 G; 20 T;

Query Match 49.3%; Score 13.8; DB 1; Length 67;
Best Local Similarity 88.2%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 accaatlttgaccaca 21
DB 51 ACCAATTTTGCCACAA 67

RESULT 15

V76550
ID V76550 standard; DNA; 71 BP.
AC V76550;
DT 16-MAR-1999 (first entry)
DE Staphylococcus aureus contig SRQ ID #2239.
KM Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
OS Staphylococcus aureus.
PN EP-786519-A2.
PD 30-JUL-1997.
PE 07-JAN-1997; 100117.
PR 05-JAN-1996; US-009861.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
PI Rosen CA;
DR WPI; 97-374922/35.
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
PS Claim 1; Page 2185; 3271pp; English.
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.
SQ Sequence 71 BP; 21 A; 21 C; 14 G; 15 T;

Query Match 49.3%; Score 13.8; DB 1; Length 71;
Best Local Similarity 72.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 cttagcaatttgaccacata 25
DB 13 CTTGACCAATAATATGATGACACATA 37

Search completed: June 4, 2000, 16:24:08
Job time: 28904 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:42 ; Search time 244.64 Seconds
(without alignments)
9.564 Million cell updates/sec

Title: US-09-164-714-12

Perfect score: 18

Sequence: 1 ggcacaaaccacccctag 18

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:*
- 4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:*
- 5: /cgn2_6/ptodata/2/1na/6.COMB.seq:*
- 6: /cgn2_6/ptodata/2/1na/PCFUS.COMB.seq:*
- 7: /cgn2_6/ptodata/2/1na/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13.8	76.7	25	1	US-08-471-206-10
2	13.8	76.7	25	2	US-08-281-025D-2
3	13.8	76.7	30	1	US-08-471-206-27
4	13.8	76.7	31	1	US-08-471-206-20
5	13.8	76.7	31	1	US-08-471-206-21
6	13.8	76.7	31	2	US-08-281-025D-14
7	13.8	76.7	31	2	US-08-281-025D-15
8	13.8	76.7	50	1	US-08-471-206-26
9	12.2	67.8	50	2	US-08-281-025D-18
10	12.2	67.8	50	2	US-08-281-025D-19
11	12.2	67.8	57	7	5221624-13
12	12.2	67.8	57	7	5221624-15
13	12.2	67.8	57	7	5221624-17
14	12.2	67.8	57	7	5221624-19
15	12.2	67.8	63	7	5221624-23
16	12.2	67.8	64	3	US-08-244-434-5
17	12.2	67.8	66	7	5221624-25
18	12.2	67.8	69	2	US-08-244-492A-8
19	12.2	67.8	69	2	US-08-709-915-10
20	12.2	67.8	75	6	PCT-US94-09653A-32
21	12.2	67.8	84	3	US-08-369-829A-15
22	12.2	67.8	84	4	US-08-586-676E-20
23	12.2	67.8	84	6	PCT-US94-09653A-33
24	12.2	67.8	94	1	US-08-315-695-26
25	11.8	65.6	23	4	US-08-793-410-14
26	11.8	65.6	89	2	US-08-290-592E-38
27	11.8	65.6	89	6	PCT-US96-09448-38

28	11.6	64.4	21	3	US-08-468-819-35	Sequence 35, Appl
29	11.6	64.4	21	3	US-08-468-819-59	Sequence 59, Appl
30	11.6	64.4	42	2	US-08-464-136-46	Sequence 46, Appl
31	11.6	64.4	42	2	US-08-440-787A-45	Sequence 45, Appl
32	11.6	64.4	42	3	US-08-349-131-46	Sequence 46, Appl
33	11.6	64.4	42	5	US-08-470-297A-46	Sequence 45, Appl
34	11.6	64.4	42	6	PCT-US91-07141-45	Sequence 45, Appl
35	11.6	64.4	42	6	PCT-US91-07149-46	Sequence 46, Appl
36	11.4	63.3	20	4	US-09-289-368-81	Sequence 81, Appl
37	11.4	63.3	23	1	US-08-433-126A-135	Sequence 135, App
38	11.4	63.3	23	2	US-08-433-126A-135	Sequence 135, App
39	11.4	63.3	23	6	PCT-US96-06059-135	Sequence 135, App
40	11.4	63.3	25	5	US-09-071-101-6	Sequence 6, Appl
41	11.4	63.3	30	1	US-08-433-126A-136	Sequence 136, App
42	11.4	63.3	30	2	US-08-433-126A-136	Sequence 136, App
43	11.4	63.3	30	6	PCT-US96-06059-136	Sequence 136, App
44	11.4	63.3	31	1	US-08-433-126A-137	Sequence 137, App
45	11.4	63.3	31	2	US-08-433-126A-137	Sequence 137, App

ALIGNMENTS

RESULT 1
US-08-471-206-10/c
Sequence 10, Application US/08471206
Patent No. 5683893
GENERAL INFORMATION:
APPLICANT: OHI, HIDEYUKI
APPLICANT: MIURA, MASAMI
APPLICANT: UNO, SHUSEI
APPLICANT: CHUGANITSU, MASAKO
APPLICANT: HIRUMATSU, RYUJI
APPLICANT: OHMURA, TAKAO
TITLE OF INVENTION: MUTANT AOX2 PROMOTER, VECTOR CARRYING
TITLE OF INVENTION: THE SAME, TRANSFORMANT, AND PRODUCTION OF HETEROLOGOUS
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPHEE & SEAS
STREET: 2100 PENNSYLVANIA AVE., NW
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08471,206
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/143,158
FILING DATE: 29-OCT-1993
APPLICATION NUMBER: JP 215306/1993
FILING DATE: 06-AUG-1993
PRIOR APPLICATION DATA: JP 293315/1992
FILING DATE: 30-OCT-1992
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-471-206-10

Query Match 76.7%; Score 13.8; DB 1; Length 25;
Best Local Similarity 88.2%; Pred. No. 52;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaacacgccta 17
|||||
DB 19 GCGACAAAATAGCCTA 3

RESULT 2

US-08-281-025D-2/c
; Sequence 2, Application US/08281025D
; Patent No. 5707827

; GENERAL INFORMATION:

; APPLICANT: OHI, Hideyuki

; APPLICANT: Mura, Masami

; APPLICANT: Hiramatsu, Ryuji

; APPLICANT: Ohmura, Takao

; TITLE OF INVENTION: MUTANT AOX2 PROMOTER, VECTOR CARRYING

; TITLE OF INVENTION: SAME, TRANSFORMANT, AND PRODUCTION OF HETEROLOGOUS PROTEIN

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

; STREET: 2100 Pennsylvania Avenue, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/281.025D

; FILING DATE: 27-JUL-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 185003/1993

; FILING DATE: 27-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Nakamura, Dean H.

; REGISTRATION NUMBER: 33,981

; REFERENCE/DOCKET NUMBER: O-35978

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)293-7060

; TELEFAX: (202)293-7860

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 25 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: other nucleic acid

; DESCRIPTION: /desc = "Plasmid DNA"

US-08-281-025D-2

Query Match 76.7%; Score 13.8; DB 2; Length 25;
Best Local Similarity 88.2%; Pred. No. 52;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaacacgccta 17
|||||
DB 19 GCGACAAAATAGCCTA 3

RESULT 3

US-08-471-206-27/c
; Sequence 27, Application US/08471206

; Patent No. 5683893

; GENERAL INFORMATION:

; APPLICANT: OHI, HIDEYUKI

; APPLICANT: MURA, MASAMI

; APPLICANT: UNO, SHUSEI

; APPLICANT: CHUGANJUT, MASAKO

; APPLICANT: HIRUMATSU, RYUJI

; APPLICANT: OHMURA, TAKAO

; TITLE OF INVENTION: MUTANT AOX2 PROMOTER, VECTOR CARRYING

; TITLE OF INVENTION: THE SAME, TRANSFORMANT, AND PRODUCTION OF HETEROLOGOUS

; TITLE OF INVENTION: PROTEIN

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

; STREET: 2100 PENNSYLVANIA AVE., NW

; CITY: WASHINGTON

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US 08/143,158

; FILING DATE: 29-OCT-1993

; APPLICATION NUMBER: JP 215306/1993

; FILING DATE: 06-AUG-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 293315/1992

; FILING DATE: 30-OCT-1992

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 293-7060

; TELEFAX: (202) 293-7860

; TELEX: 6491103

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 30 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-471-206-27

Query Match 76.7%; Score 13.8; DB 1; Length 30;
Best Local Similarity 88.2%; Pred. No. 53;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaacacgccta 17
|||||
DB 23 GCGACAAAATAGCCTA 7

RESULT 4

US-08-471-206-20/c
; Sequence 20, Application US/08471206
; Patent No. 5683893

; GENERAL INFORMATION:

; APPLICANT: OHI, HIDEYUKI

; APPLICANT: MURA, MASAMI

; APPLICANT: UNO, SHUSEI

; APPLICANT: CHUGANJUT, MASAKO

; APPLICANT: HIRUMATSU, RYUJI

; APPLICANT: OHMURA, TAKAO

; TITLE OF INVENTION: MUTANT AOX2 PROMOTER, VECTOR CARRYING

; TITLE OF INVENTION: THE SAME, TRANSFORMANT, AND PRODUCTION OF HETEROLOGOUS

; TITLE OF INVENTION: PROTEIN

NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 PENNSYLVANIA AVE., NW
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,206
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/143,158
FILING DATE: 29-OCT-1993
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: Synthetic DNA
US-08-471-206-20

Query Match 76.7%; Score 13.8; DB 1; Length 31;
Best Local Similarity 88.2%; Pred. No. 53;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaaccagccta 17
||||||| |||||
Db 24 GCGACAAAATAGCCTA 8

RESULT 5
US-08-471-206-21
Sequence 21, Application US/08471206
Patent No. 5683893
GENERAL INFORMATION:
APPLICANT: OHI, HIDEYUKI
APPLICANT: MIURA, MASAMI
APPLICANT: UNO, SHUSEI
APPLICANT: CHUGANJI, MASAKO
APPLICANT: HIRUMATSU, RYUJI
APPLICANT: OHMURA, TAKAO
TITLE OF INVENTION: MUTANT AOX2 PROMOTER, VECTOR CARRYING
TITLE OF INVENTION: THE SAME, TRANSFORMANT, AND PRODUCTION OF HETEROLOGOUS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 PENNSYLVANIA AVE., NW
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,206
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/143,158
FILING DATE: 29-OCT-1993
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: Synthetic DNA
US-08-471-206-21

Query Match 76.7%; Score 13.8; DB 1; Length 31;
Best Local Similarity 88.2%; Pred. No. 53;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gcgacaaaccagccta 17
||||||| |||||
Db 12 GCGACAAAATAGCCTA 28

RESULT 6
US-08-281-025D-14/C
Sequence 14, Application US/08281025D
Patent No. 5707827
GENERAL INFORMATION:
APPLICANT: Ohi, Hideyuki
APPLICANT: Miura, Masami
APPLICANT: Hiramatsu, Ryuji
APPLICANT: Ohmura, Takao
TITLE OF INVENTION: MUTANT AOX2 PROMOTER, VECTOR CARRYING
TITLE OF INVENTION: SAME, TRANSFORMANT, AND PRODUCTION OF HETEROLOGOUS PROTEIN
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/281,025D
FILING DATE: 27-JUL-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 185003/1993
FILING DATE: 27-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Dean H.
REGISTRATION NUMBER: 33,981

REFERENCE/DOCKET NUMBER: Q-35978
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Primer"
US-08-281-025D-14

Query Match 76.7%; Score 13.8; DB 2; Length 31;
Best Local Similarity 88.2%; Pred. No. 53;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gcgacaaaccagccta 17
||||||| |||||
Db 24 GCGACAAATAAGCCTA 8

RESULT 7
US-08-281-025D-15
Sequence 15, Application US/08281025D
Patent No. 5707827
GENERAL INFORMATION:
APPLICANT: Ohl, Hideyuki
APPLICANT: Miura, Masami
APPLICANT: Hiramatsu, Ryuji
APPLICANT: Ohmura, Takao
TITLE OF INVENTION: MUTANT AOX2 PROMOTER, VECTOR CARRYING
TITLE OF INVENTION: SAME, TRANSFORMANT, AND PRODUCTION OF HETEROLOGOUS PROTEIN
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MIOW, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/281,025D
FILING DATE: 27-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 185003/1993
FILING DATE: 27-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Dean H.
REGISTRATION NUMBER: 33,981
REFERENCE/DOCKET NUMBER: Q-35978
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Primer"
US-08-281-025D-15

Query Match 76.7%; Score 13.8; DB 2; Length 31;
Best Local Similarity 88.2%; Pred. No. 53;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 gcgacaaaccagccta 17
||||||| |||||
Db 12 GCGACAAATAAGCCTA 28

RESULT 8
US-08-471-206-26/c
Sequence 26, Application US/08471206
Patent No. 5683893
GENERAL INFORMATION:
APPLICANT: OHI, HIDEYUKI
APPLICANT: MIURA, MASAMI
APPLICANT: UNO, SHUSEI
APPLICANT: CHOGANUJI, MASAKO
APPLICANT: HIRUMATSU, RYUJI
APPLICANT: OHMURA, TAKAO
TITLE OF INVENTION: MUTANT AOX2 PROMOTER, VECTOR CARRYING
TITLE OF INVENTION: THE SAME, TRANSFORMANT, AND PRODUCTION OF HETEROLOGOUS
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MIOW, ZINN, MACPEAK & SEAS
STREET: 2100 PENNSYLVANIA AVE., NW
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,206
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/143,158
FILING DATE: 29-OCT-1993
APPLICATION NUMBER: JP 215306/1993
FILING DATE: 06-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 293315/1992
FILING DATE: 30-OCT-1992
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-471-206-26

Query Match 76.7%; Score 13.8; DB 1; Length 50;
Best Local Similarity 88.2%; Pred. No. 56;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gcgacaaaccagccta 17
||||||| |||||
Db 41 GCGACAAATAAGCCTA 25

RESULT 9
US-08-281-025D-18/c

```
; Sequence 18, Application US/08281025D
; Patent No. 5707827
; GENERAL INFORMATION:
; APPLICANT: Ohl, Hideyuki
; APPLICANT: Miura, Masami
; APPLICANT: Hirata, Ryuji
; APPLICANT: Ohmura, Takao
; TITLE OF INVENTION: MUTANT AOX2 PROMOTER, VECTOR CARRYING
; TITLE OF INVENTION: SAME, TRANSFORMANT, AND PRODUCTION OF HETEROLOGOUS PROTEIN
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/281,025D
; FILING DATE: 27-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 185003/1993
; FILING DATE: 27-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Dean H.
; REGISTRATION NUMBER: 33,981
; REFERENCE/DOCKET NUMBER: Q-35978
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Primer"
; US-08-281-025D-18

Query Match          67.8%; Score 12.2; DB 2; Length 50;
Best Local Similarity 82.4%; Pred. No. 3.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gcgacaaacagccta 17
    |||||
DB 41 GCGACAAAATGAGCCCA 25

RESULT 10
US-08-281-025D-19/c
; Sequence 19, Application US/08281025D
; Patent No. 5707827
; GENERAL INFORMATION:
; APPLICANT: Ohl, Hideyuki
; APPLICANT: Miura, Masami
; APPLICANT: Hirata, Ryuji
; APPLICANT: Ohmura, Takao
; TITLE OF INVENTION: MUTANT AOX2 PROMOTER, VECTOR CARRYING
; TITLE OF INVENTION: SAME, TRANSFORMANT, AND PRODUCTION OF HETEROLOGOUS PROTEIN
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
```

```
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/281,025D
; FILING DATE: 27-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 185003/1993
; FILING DATE: 27-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Dean H.
; REGISTRATION NUMBER: 33,981
; REFERENCE/DOCKET NUMBER: Q-35978
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Primer"
; US-08-281-025D-19

Query Match          67.8%; Score 12.2; DB 2; Length 50;
Best Local Similarity 82.4%; Pred. No. 3.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gcgacaaacagccta 17
    |||||
DB 41 GAGACAAAATAGCCCTA 25
```

```
RESULT 11
5221624-13/c
; Patent No. 5221624
; APPLICANT: BLAIR, LINDLEY C.; KODURI, JAR-HOW; WEICKMANN,
; JOACHIM J.
; TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
; (LYS46, ASP113, ASP137) THAUMATIN I
; NUMBER OF SEQUENCES: 31
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/407,416
; FILING DATE: 14-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 268,702
; FILING DATE: 08-NOV-1988
; SEQ ID NO:13
; LENGTH: 57
5221624-13

Query Match          67.8%; Score 12.2; DB 7; Length 57;
Best Local Similarity 82.4%; Pred. No. 3.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gcgacaaacagccta 17
    |||||
DB 44 GCTGCAAAACAGCCCA 28

RESULT 12
5221624-15/c
; Patent No. 5221624
; APPLICANT: BLAIR, LINDLEY C.; KODURI, JAR-HOW; WEICKMANN,
```

; JOACHIM J.
; TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
; (LYS46, ASP113, ASP137) THAUMATIN I
; NUMBER OF SEQUENCES: 31
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/407,416
; FILING DATE: 14-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 268,702
; FILING DATE: 08-NOV-1988
; SEQ ID NO:15:
; LENGTH: 57
5221624-15

Query Match 67.8%; Score 12.2; DB 7; Length 57;
Best Local Similarity 82.4%; Pred. No. 3.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gcgacaaaaccagccta 17
|| |||||
DB 44 GCTGCAAAACCGCCAA 28

RESULT 13
5221624-17/c
; Patent No. 5221624
; APPLICANT: BLAIR, LINDLEY C.; KODURI, JAR-HOW; WEICKMANN,
; JOACHIM J.
; TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
; (LYS46, ASP113, ASP137) THAUMATIN I
; NUMBER OF SEQUENCES: 31
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/407,416
; FILING DATE: 14-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 268,702
; FILING DATE: 08-NOV-1988
; SEQ ID NO:17:
; LENGTH: 57
5221624-17

Query Match 67.8%; Score 12.2; DB 7; Length 57;
Best Local Similarity 82.4%; Pred. No. 3.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gcgacaaaaccagccta 17
|| |||||
DB 44 GCTGCAAAACCGCCAA 28

RESULT 14
5221624-19/c
; Patent No. 5221624
; APPLICANT: BLAIR, LINDLEY C.; KODURI, JAR-HOW; WEICKMANN,
; JOACHIM J.
; TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
; (LYS46, ASP113, ASP137) THAUMATIN I
; NUMBER OF SEQUENCES: 31
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/407,416
; FILING DATE: 14-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 268,702
; FILING DATE: 08-NOV-1988
; SEQ ID NO:19:
; LENGTH: 57
5221624-19

Query Match 67.8%; Score 12.2; DB 7; Length 57;
Best Local Similarity 82.4%; Pred. No. 3.8e+02;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 gcgacaaaaccagccta 17
|| |||||
DB 44 GCTGCAAAACCGCCAA 28

RESULT 15
5221624-23/c
; Patent No. 5221624
; APPLICANT: BLAIR, LINDLEY C.; KODURI, JAR-HOW; WEICKMANN,
; JOACHIM J.
; TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
; (LYS46, ASP113, ASP137) THAUMATIN I
; NUMBER OF SEQUENCES: 31
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/407,416
; FILING DATE: 14-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 268,702
; FILING DATE: 08-NOV-1988
; SEQ ID NO:23:
; LENGTH: 63
5221624-23

Query Match 67.8%; Score 12.2; DB 7; Length 63;
Best Local Similarity 82.4%; Pred. No. 3.9e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gcgacaaaaccagccta 17
|| |||||
DB 50 GCTGCAAAACCGCCAA 34

Search completed: June 4, 2000, 16:09:43
Job time: 28062 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:10 ; Search time 322.35 Seconds
(without alignments)
13.971 Million cell updates/sec

Title: US-09-164-714-13

Perfect score: 18
Sequence: 1 ggtgtgtgtgtgtgtt 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues
Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	13.8	76.7	36	1	Q55562	Sequence of synthe
2	13.8	76.7	36	1	Q55561	Sequence of contro
3	13.8	76.7	36	1	Q55569	Sequence of synthe
4	13.8	76.7	36	1	Q55574	Sequence of synthe
5	13.8	76.7	36	1	Q67225	Triple helix-forml
6	13.8	76.7	36	1	Q67229	Triple helix-forml
7	13.8	76.7	36	1	Q67232	Triple helix-forml
8	13.8	76.7	36	1	Q79205	Guanosine rich oli
9	13.8	76.7	36	1	Q79207	Guanosine rich oli
10	13.8	76.7	36	1	T51623	Viral integrase in
11	13.8	76.7	36	1	T51625	Viral integrase in
12	13.8	76.7	42	1	X27667	DNA template 5'-en
13	13.4	74.4	24	1	V61397	Primer Mu-Bio used
14	13.4	74.4	27	1	V07668	Sequence of Mu-Bio
15	13.4	74.4	27	1	X32517	Human gene signatu
16	13.4	74.4	77	1	T20428	G-quartet oligonuc
17	13.4	73.3	20	1	T97948	Synthetic oligonuc
18	13.2	73.3	24	1	N80521	Guanosine rich oli
19	13.2	73.3	24	1	Q79216	Viral integrase in
20	13.2	73.3	24	1	T51634	Triplex forming ol
21	13.2	73.3	35	1	Q33260	S hylicus lipase ge
22	13.2	73.3	41	1	T27080	Brassica sp. polym
23	13.2	73.3	41	1	V50702	E. coli tryptophan
24	13.2	73.3	45	1	Q69019	Transmembrane olig
25	13.2	73.3	45	1	Q69017	Human cyclin E gen
26	13.2	73.3	46	1	Q94193	Human IL-5 gene pr
27	13.2	73.3	67	1	V13178	CRCA-1 coding sequ
28	13.2	73.3	72	1	X25876	Human ballelic po
29	13.2	73.3	78	1	X27805	MuH1 gene specific
30	12.8	71.1	18	1	X09673	Thyrotropin recept
31	12.8	71.1	22	1	X32038	Sequence binding t
32	12.8	71.1	30	1	V02229	APPpar, targeted
33	12.8	71.1	35	1	Q04993	
34	12.8	71.1	35	1	Q36271	

35	12.8	71.1	35	1	Q36272	APPsant, targete
36	12.8	71.1	38	1	Q05015	Sequence binding t
37	12.8	71.1	38	1	Q36337	HSVORLpar, target
38	12.8	71.1	38	1	Q36338	HSVORLpar, target
39	12.8	71.1	53	1	X23820	Bacillus sp. GUS g
40	12.8	71.1	66	1	V77825	Staphylococcus aur
41	12.8	71.1	76	1	Q03555	DNA sequence encod
42	12.8	71.1	76	1	N92819	Modified type of s
43	12.8	71.1	77	1	Q03553	DNA sequence of 1m
44	12.8	71.1	77	1	N92360	Gene for modified
45	12.8	71.1	79	1	N81283	Sequence encoding

ALIGNMENTS

RESULT	ID	Score	Match	Length	DB	ID	Description
055562	Q55562	standard	DNA	36	BP.		
AC	Q55562	17-AUG-1994	(first entry)				
DE	Sequence of synthetic triplex forming oligo (TFO) Seq ID no. 2.						
KW	Triplex forming oligonucleotide; TFO; sequence specific binding; ss.						
OS	Synthetic.						
FH	Key	Location/Qualifiers					
FT	misc_feature	36					
FT		/*tag= a					
PN	W09404550-A.	/label= G-Chol.					
PD	03-MAR-1994.						
PF	17-AUG-1993; U07743.						
PR	21-AUG-1992; US-934065.						
PR	23-APR-1993; US-053040.						
PA	(BAYU) BAYLOR COLLEGE MEDICINE.						
PI	(TRIP-) TRIPLEX PHARM CORP.						
PI	Hogan ME, Jayaraman K, Vu H, Zendegeu J;						
DR	WPI: 94-083097/10						
PT	New method of binding synthetic triplex forming oligonucleotide -						
PT	in which the nucleotide is modified with a lipophilic cpd. is						
PT	useful in treatment of cell proliferative states and viral						
PS	infections						
PS	Claim 8: Page 24; 86pp: English.						
CC	Triplex forming oligos (TFOs) bind to DNA in a site selective manner.						
CC	The biological effect of a TFO is potentiated by modification with						
CC	lipophilic cpds., selected from cholesterol, vitamin E and 1,2-di-O-						
CC	hexadecyl-3-glycerol. TFO-linker-cholesterol is used for the						
CC	treatment of cell proliferative states (breast, lung and cervical						
CC	cancers) and in infections by viruses (Herpes simplex virus type 2						
CC	and human immunodeficiency virus (HIV)).						
CC	Sequence 36 BP; 0 A; 0 C; 24 G; 12 T;						
SQ							
Query Match	76.7%;	Score 13.8;	DB 1;	Length 36;			
Best Local Similarity	88.2%;	Pred. No. 4.4e+02;					
Matches 15;	Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;		
OY	1 ggtgtgtgtgtgtgtt 17						
Db	9 GGTGTGTGTGTGTGT 25						
RESULT	2						
Q55561	Q55561	standard	DNA	36	BP.		
ID	Q55561	17-AUG-1994	(first entry)				
AC	Q55561	Sequence of control for synthetic triplex forming oligo (TFO)					
DE	Seq ID no. 2.						
KW	Triplex forming oligonucleotide; TFO; sequence specific binding; ss.						
OS	Synthetic.						
FH	Key	Location/Qualifiers					
FT	misc_feature	36					
FT		/*tag= a					

FT /label= G-NH2.
 PN W09404550-A.
 PD 03-MAR-1994.
 PE 17-AUG-1993; U07743.
 PR 21-AUG-1992; US-934065.
 PR 23-APR-1993; US-053040.
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 PI (TRIP-) TRIPLEX PHARM CORP.
 PI Hogan ME, Jayaraman K, Vu H, Zendequi J;
 DR WPI; 94-083097/10.
 PT New method of binding synthetic triplex forming oligonucleotide -
 PT in which the nucleotide is modified with a lipophilic cpd. is
 PT useful in treatment of cell proliferative states and viral
 PT infections
 PS Example; Page 24; 86pp; English.
 CC Triplex forming oligos (TFOs) bind to DNA in a site selective manner.
 CC The biological effect of a TFO is potentiated by modification with
 CC lipophilic cpds.. selected from cholesterol, vitamin E and 1,2-di-o-
 CC hexadecyl-3-glyceryl. TFO-linker-cholesterol is used for the
 CC treatment of cell proliferative states (breast, lung and cervical
 CC cancers) and in infections by viruses (Herpes simplex virus type 2
 CC and human immunodeficiency virus (HIV)).
 SQ Sequence 36 BP; 0 A; 0 C; 24 G; 12 T;

Query Match 76.7%; Score 13.8; DB 1; Length 36;
 Best Local Similarity 88.2%; Pred. No. 4.4e+02;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgtt 17
 |||||
 DB 9 GGTGTTGCTGCTGCTT 25

RESULT 3
 Q55569
 ID Q55569 standard; DNA; 36 BP.
 AC Q55569;
 DE 17-AUG-1994 (first entry)
 DE Sequence of synthetic triplex forming oligo (TFO) B-106-85,
 DE anti-HSV 2 TFO
 KM Triplex forming oligonucleotide; TFO; sequence specific binding; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 36
 FT /*tag= a
 FT /label= G-cholesterol
 PN W09404550-A.
 PD 03-MAR-1994.
 PE 17-AUG-1993; U07743.
 PR 21-AUG-1992; US-934065.
 PR 23-APR-1993; US-053040.
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 PI (TRIP-) TRIPLEX PHARM CORP.
 PI Hogan ME, Jayaraman K, Vu H, Zendequi J;
 DR WPI; 94-083097/10.
 PT New method of binding synthetic triplex forming oligonucleotide -
 PT in which the nucleotide is modified with a lipophilic cpd. is
 PT useful in treatment of cell proliferative states and viral
 PT infections
 PS Example; Page 40; 86pp; English.
 CC Triplex forming oligos (TFOs) bind to DNA in a site selective manner.
 CC The biological effect of a TFO is potentiated by modification with
 CC lipophilic cpds.. selected from cholesterol, vitamin E and 1,2-di-o-
 CC hexadecyl-3-glyceryl. TFO-linker-cholesterol is used for the
 CC treatment of cell proliferative states (breast, lung and cervical
 CC cancers) and in infections by viruses (Herpes simplex virus type 2
 CC and human immunodeficiency virus (HIV)).
 SQ Sequence 36 BP; 0 A; 0 C; 24 G; 12 T;

Query Match 76.7%; Score 13.8; DB 1; Length 36;
 Best Local Similarity 88.2%; Pred. No. 4.4e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 ggtgtgtgtgtgtt 17
 |||||
 DB 9 GGTGTTGCTGCTGCTT 25

RESULT 4
 Q55574
 ID Q55574 standard; DNA; 36 BP.
 AC Q55574;
 DE 17-AUG-1994 (first entry)
 DE Sequence of synthetic triplex forming oligo (TFO)
 DE B-106-89, anti-HSV2 TFO
 KM Triplex forming oligonucleotide; TFO; sequence specific binding; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 1
 FT /*tag= a
 FT /label= Cholesterol-G
 FT misc_feature 36
 FT /*tag= b
 FT /label= G-propanolamine
 PN W09404550-A.
 PD 03-MAR-1994.
 PE 17-AUG-1993; U07743.
 PR 21-AUG-1992; US-934065.
 PR 23-APR-1993; US-053040.
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 PI (TRIP-) TRIPLEX PHARM CORP.
 PI Hogan ME, Jayaraman K, Vu H, Zendequi J;
 DR WPI; 94-083097/10.
 PT New method of binding synthetic triplex forming oligonucleotide -
 PT in which the nucleotide is modified with a lipophilic cpd. is
 PT useful in treatment of cell proliferative states and viral
 PT infections
 PS Example; Page 41; 86pp; English.
 CC Triplex forming oligos (TFOs) bind to DNA in a site selective manner.
 CC The biological effect of a TFO is potentiated by modification with
 CC lipophilic cpds.. selected from cholesterol, vitamin E and 1,2-di-o-
 CC hexadecyl-3-glyceryl. TFO-linker-cholesterol is used for the
 CC treatment of cell proliferative states (breast, lung and cervical
 CC cancers) and in infections by viruses (Herpes simplex virus type 2
 CC and human immunodeficiency virus (HIV)). Q55573-75 were synthesised
 CC using the cholesterol phosphoramidite, synthesis of
 CC 2-cyanoethyl-N,N'-diisopropyl-1-O-(4,4'-dimethoxytrityl)-2'-N-
 CC ((N'-cholesteryl)oxycarbonyl)glycidylglycidylaminopropaneglycyl
 CC phosphoramidite.
 SQ Sequence 36 BP; 0 A; 0 C; 24 G; 12 T;

Query Match 76.7%; Score 13.8; DB 1; Length 36;
 Best Local Similarity 88.2%; Pred. No. 4.4e+02;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtt 17
 |||||
 DB 9 GGTGTTGCTGCTGCTT 25

RESULT 5
 Q67225
 ID Q67225 standard; DNA; 36 BP.
 AC Q67225;
 DE 22-SEP-1994 (first entry)
 DE Triplex helix-forming oligonucleotide.
 DE Triple helix; oligonucleotide; TFO; lipophilic; cholesterol; Vitamin E;
 KM glyceryl; anticancer; antiviral; herpes simplex; HIV; human;
 KM immunodeficiency virus.
 OS Synthetic.
 PN W09404550-A.
 PD 03-MAR-1994.
 PE 17-AUG-1993; U07743.

PR 21-AUG-1992; US-934065.
PR 23-APR-1993; US-053040.
PA (BAYU) BAYLOR COLLEGE MEDICINE.
PA (TRIP-) TRIPLEX PHARM CORP.
PI Hogan ME, Jayaraman K, Vu H, Zengdegi J;
DR WPI: 94-083097/10.
PT New method of binding synthetic triplex forming oligonucleotide -
PT in which the nucleotide is modified with a lipophilic cpd; is
PT useful in treatment of cell proliferative states and viral
PT infections
PS Claim 8; Page 24; 86pp; English.
CC The invention relates to a method of enhancing sequence specific
CC binding of a synthetic triplex-forming oligonucleotide (TFO)
CC involving the step of contacting the TFO with a cell. The TFO comprises
CC a nucleotide sequence of at least 20 nucleotides including a G and a T,
CC is chemically modified with a lipophilic compound and is capable of
CC binding to a DNA duplex target to form a triplex helix. The lipophilic
CC compound is preferably cholesterol, vitamin E or 1,2-di-O-hexadecyl
CC 3-glycerol and is joined to the nucleotide via a linker.
CC The TFO is useful medically for treating cell proliferative
CC states such as breast, lung and cervical cancers, and for treating

Query Match 76.7%; Score 13.8; DB 1; Length 36;
Best Local Similarity 88.2%; Pred. No. 4.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgt 17
||||||| ||| ||
Db 9 GGTGTGTGTGTGTGT 25

RESULT 6
067229
ID 067229 standard; DNA; 36 BP.
AC 067229;
DT 22-SEP-1994 (first entry)
DE Triplex helix-forming oligonucleotide.
KW Triplex helix; oligonucleotide; TFO; lipophilic; cholesterol; Vitamin E;
KW glyceryl; anticancer; antiviral; herpes simplex; HIV; human;
KM immunodeficiency virus.
OS Synthetic.
PN WO9404550-A.
PD 03-MAR-1994.
PF 17-AUG-1993; U07743.
PR 21-AUG-1992; US-934065.
PR 23-APR-1993; US-053040.
PA (BAYU) BAYLOR COLLEGE MEDICINE.
PA (TRIP-) TRIPLEX PHARM CORP.
PI Hogan ME, Jayaraman K, Vu H, Zengdegi J;
DR WPI: 94-083097/10.
PT New method of binding synthetic triplex forming oligonucleotide -
PT in which the nucleotide is modified with a lipophilic cpd; is
PT useful in treatment of cell proliferative states and viral
PT infections
PS Claim 25; Page 40; 86pp; English.
CC The invention relates to a method of enhancing sequence specific
CC binding of a synthetic triplex-forming oligonucleotide (TFO)
CC involving the step of contacting the TFO with a cell. The TFO comprises
CC a nucleotide sequence of at least 20 nucleotides including a G and a T,
CC is chemically modified with a lipophilic compound and is capable of
CC binding to a DNA duplex target to form a triplex helix. The lipophilic
CC compound is preferably cholesterol, vitamin E or 1,2-di-O-hexadecyl
CC 3-glycerol and is joined to the nucleotide via a linker.
CC The TFO is useful medically for treating cell proliferative
CC states such as breast, lung and cervical cancers, and for treating

CC viral infections such as caused by Herpes simplex type 2 or HIV. The
CC biological effect of the TFO is potentiated by modification with the
CC lipophilic compound.
CC The present sequence is one of 9 specific nucleotide sequences disclosed
CC for use in forming the TFO (067224 - 067232). The TFO is formed by
CC attaching cholesterol to the 3' end. This TFO is active against
CC HSV-2.
SQ Sequence 36 BP; 0 A; 0 C; 24 G; 12 T;

Query Match 76.7%; Score 13.8; DB 1; Length 36;
Best Local Similarity 88.2%; Pred. No. 4.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgt 17
||||||| ||| ||
Db 9 GGTGTGTGTGTGTGT 25

RESULT 7
067232
ID 067232 standard; DNA; 36 BP.
AC 067232;
DT 22-SEP-1994 (first entry)
DE Triplex helix-forming oligonucleotide.
KW Triplex helix; oligonucleotide; TFO; lipophilic; cholesterol; Vitamin E;
KW glyceryl; anticancer; antiviral; herpes simplex; HIV; human;
KM immunodeficiency virus.
OS Synthetic.
PN WO9404550-A.
PD 03-MAR-1994.
PF 17-AUG-1993; U07743.
PR 21-AUG-1992; US-934065.
PR 23-APR-1993; US-053040.
PA (BAYU) BAYLOR COLLEGE MEDICINE.
PA (TRIP-) TRIPLEX PHARM CORP.
PI Hogan ME, Jayaraman K, Vu H, Zengdegi J;
DR WPI: 94-083097/10.
PT New method of binding synthetic triplex forming oligonucleotide -
PT in which the nucleotide is modified with a lipophilic cpd; is
PT useful in treatment of cell proliferative states and viral
PT infections
PS Claim 25; Page 41; 86pp; English.
CC The invention relates to a method of enhancing sequence specific
CC binding of a synthetic triplex-forming oligonucleotide (TFO)
CC involving the step of contacting the TFO with a cell. The TFO comprises
CC a nucleotide sequence of at least 20 nucleotides including a G and a T,
CC is chemically modified with a lipophilic compound and is capable of
CC binding to a DNA duplex target to form a triplex helix. The lipophilic
CC compound is preferably cholesterol, vitamin E or 1,2-di-O-hexadecyl
CC 3-glycerol and is joined to the nucleotide via a linker.
CC The TFO is useful medically for treating cell proliferative
CC states such as breast, lung and cervical cancers, and for treating
CC viral infections such as caused by Herpes simplex type 2 or HIV. The
CC biological effect of the TFO is potentiated by modification with the
CC lipophilic compound.
CC The present sequence is one of 9 specific nucleotide sequences disclosed
CC for use in forming the TFO (067224 - 067232). The TFO is formed by
CC attaching cholesterol to the 5' end and propanolamine to the 3' end.
CC This TFO is active against HSV-2.
SQ Sequence 36 BP; 0 A; 0 C; 24 G; 12 T;

Query Match 76.7%; Score 13.8; DB 1; Length 36;
Best Local Similarity 88.2%; Pred. No. 4.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgt 17
||||||| ||| ||
Db 9 GGTGTGTGTGTGTGT 25

RESULT 8

```
Q79205
ID Q79205 standard; DNA; 36 BP.
AC Q79205;
DE 17-JUL-1995 (first entry)
KW Guanostine rich oligonucleotide used to treat viral infection.
KW Guanostine; tetrad; inhibition; replication; virus; treatment;
KW therapy; infection; herpes simplex virus; human papilloma virus;
KW Epstein-Barr virus; HIV; adenovirus; respiratory syncytial virus;
KW hepatitis B virus; human cytomegalovirus; ss.
OS Synthetic.
FH Key
FT misc_feature
FT 36 Location/Qualifiers
FT /*tag= a
FT /mod_base=
FT /note= "Propanolamine group attached to this base."
PN MO9425037-A.
PD 10-NOV-1994.
PE 25-APR-1994; U04529.
PR 23-APR-1993; US-053027.
PR 28-OCT-1993; US-145704.
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
PA (TRIP-) TRIPLEX PHARM CORP.
PI Fennewald S, Hogan ME, Ojwang JO, Rando RF, Zengdegi JG;
PI WPI: 94-357890/44.
PT Oligo-nucleotide(s) rich in guanosine which form guanosine
PT tetrads - used to treat viral infections, e.g. herpes-virus and
PT HIV
PS Claim 41; Page 46; 101pp; English.
CC The oligonucleotides (See Q79201-52) can be used to treat viral
CC infections. The oligonucleotides inhibit viral replication by
CC forming guanosine tetrads which form a stabilised 3D structure.
CC Preferred oligonucleotides contain at least 2 runs of at least 2
CC guanosine bases and may be capped at the 3' terminus with a modifier
CC selected from polyamine, poly-L-lysine, cholesterol and
CC propanolamine. They may also have a modified phosphodiester linkage
CC or be modified to contain a phosphorothioate linkage. They are used
CC to treat infections with viruses such as herpes simplex virus, human
CC papilloma virus, Epstein-Barr virus, HIV, adenovirus, respiratory
CC syncytial virus, hepatitis B virus or human cytomegalovirus.
SQ Sequence 36 BP; 0 A; 0 C; 24 G; 12 T;
```

```
Query Match 76.7%; Score 13.8; DB 1; Length 36;
Best Local Similarity 88.2%; Pred. No. 4.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ggtgtgtgtgtgtgctt 17
   |||||
DB 9 GGTGTGTGTGTGTGTT 25
```

```
RESULT 9
Q79207
ID Q79207 standard; DNA; 36 BP.
AC Q79207;
DE 17-JUL-1995 (first entry)
KW Guanostine rich oligonucleotide used to treat viral infection.
KW Guanostine; tetrad; inhibition; replication; virus; treatment;
KW therapy; infection; herpes simplex virus; human papilloma virus;
KW Epstein-Barr virus; HIV; adenovirus; respiratory syncytial virus;
KW hepatitis B virus; human cytomegalovirus; ss.
OS Synthetic.
FH Key
FT misc_feature
FT 36 Location/Qualifiers
FT /*tag= a
FT /mod_base=
FT /note= "Propanolamine group attached to this base."
FT 1. .36
FT /*tag= b
FT /note= "Phosphorothioate backbone."
PN MO9425037-A.
PD 10-NOV-1994.
PE 25-APR-1994; U04529.
```

```
PR 23-APR-1993; US-053027.
PR 28-OCT-1993; US-145704.
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
PA (TRIP-) TRIPLEX PHARM CORP.
PI Fennewald S, Hogan ME, Ojwang JO, Rando RF, Zengdegi JG;
PI WPI: 94-357890/44.
PT Oligo-nucleotide(s) rich in guanosine which form guanosine
PT tetrads - used to treat viral infections, e.g. herpes-virus and
PT HIV
PS Claim 41; Page 47; 101pp; English.
CC The oligonucleotides (See Q79201-52) can be used to treat viral
CC infections. The oligonucleotides inhibit viral replication by
CC forming guanosine tetrads which form a stabilised 3D structure.
CC Preferred oligonucleotides contain at least 2 runs of at least 2
CC guanosine bases and may be capped at the 3' terminus with a modifier
CC selected from polyamine, poly-L-lysine, cholesterol and
CC propanolamine. They may also have a modified phosphodiester linkage
CC or be modified to contain a phosphorothioate linkage. They are used
CC to treat infections with viruses such as herpes simplex virus, human
CC papilloma virus, Epstein-Barr virus, HIV, adenovirus, respiratory
CC syncytial virus, hepatitis B virus or human cytomegalovirus.
SQ Sequence 36 BP; 0 A; 0 C; 24 G; 12 T;
```

```
Query Match 76.7%; Score 13.8; DB 1; Length 36;
Best Local Similarity 88.2%; Pred. No. 4.4e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ggtgtgtgtgtgtgctt 17
   |||||
DB 9 GGTGTGTGTGTGTGTT 25
```

```
RESULT 10
TS1623
ID TS1623 standard; DNA; 36 BP.
AC TS1623;
DE 11-NOV-1997 (first entry)
KW Viral integrase inhibiting oligonucleotide.
KW Human immunodeficiency virus; HIV; Epstein Barr virus; EBV;
KW herpes simplex virus; HSV; human papilloma virus; HPV; adenovirus;
KW respiratory syncytial virus; RSV; cytomegalovirus; CMV; hepatitis B;
KW integrase inhibition; guanosine tetrad; ss.
OS Synthetic.
PN MO9703997-A1.
PD 06-FEB-1997.
PE 17-JUL-1996; U11786.
PE 23-APR-1996; US-016271.
PR 19-JUL-1995; US-001505.
PR 23-OCT-1995; US-535168.
PR 19-MAR-1996; US-013688.
PR 25-MAR-1996; US-014007.
PR 17-APR-1996; US-015714.
PA (ARON-) ARONEX PHARM INC.
PI Fennewald S, Hogan ME, Mazumder A, Ojwang JO, Pommler Y;
PI Rando RF, Zengdegi JG;
PI WPI: 97-132569/12.
PT Oligo-nucleotide(s) capable of forming guanosine tetrads - inhibit
PT viral enzyme responsible for integrating viral nucleic acid into the
PT host genome
PS Claim 3; Page 142; 245pp; English.
CC TS1619-TS1698 are oligonucleotides used to inhibit the production
CC of viruses within a host cell. The oligonucleotides may form guanosine
CC tetrads (structures formed of eight hydrogen bonds by coordination of
CC the four oxygen atoms of guanine with alkali cations believed to bind
CC to the centre of a quadruplex, and by strong stacking interactions) and
CC are used to prevent the integration of viral nucleic acid into a host
CC genome. The oligonucleotides inhibit functioning of the integrase enzyme
CC and hence prevent viral infection. Viral infections that may be treated
CC include human immunodeficiency virus (HIV), Epstein Barr virus (EBV),
CC herpes simplex virus (HSV), human papilloma virus (HPV), adenovirus,
CC respiratory syncytial virus (RSV), cytomegalovirus (CMV) and hepatitis
CC B virus (HBV), especially HIV-1 infection.
```

SQ Sequence 36 BP: 0 A; 0 C; 24 G; 12 T;
 Query Match 76.7%; Score 13.8; DB 1; Length 36;
 Best Local Similarity 88.2%; Pred. No. 4.4e+02;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 ggtgtgtgtgtgtgtgt 17
 |||||
 DB 9 GGTGTGTGTGTGTGT 25
 RESULT 11
 T51625
 ID T51625 standard; DNA: 36 BP.
 AC T51625;
 DT 11-NOV-1997 (first entry)
 DE Viral integrase inhibiting oligonucleotide.
 KW herpes simplex virus; HSV; human papilloma virus; HPV; adenovirus;
 KW respiratory syncytial virus; RSV; cytomegalovirus; CMV; hepatitis B;
 KW integrase inhibition; guanosine tetrad; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_difference 36
 FT /tag- a
 FT /note= "phosphorothioate backbone"
 PN W09703997-A1.
 PD 06-FEB-1997.
 PE 17-JUL-1996; U11786.
 PR 23-APR-1996; US-016271.
 PR 19-JUL-1995; US-001505.
 PR 23-OCT-1995; US-535188.
 PR 19-MAR-1996; US-013688.
 PR 25-MAR-1996; US-014007.
 PR 17-APR-1996; US-015714.
 PA (ARON-) ARONEX PHARM INC.
 PI Fennewald S, Hogan ME, Mazumder A, Ojwang JO, Pommler Y;
 PI Rando RF, Zendejani JG;
 PI WPI: 97-132569/12.
 PT Oligo:nucleotide(s) capable of forming guanosine tetrads - inhibit
 PT viral enzyme responsible for integrating viral nucleic acid into the
 PT host genome
 PS Claim 3; Page 143; 245pp; English.
 CC T51619-T51698 are oligonucleotides used to inhibit the production
 CC of viruses within a host cell. The oligonucleotides may form guanosine
 CC tetrads (structures formed of eight hydrogen bonds by coordination of
 CC the four oxygen atoms of guanine with alkali cations believed to bind
 CC to the centre of a quadruplex, and by strong stacking interactions) and
 CC are used to prevent the integration of viral nucleic acid into a host
 CC genome. The oligonucleotides inhibit functioning of the integrase enzyme
 CC and hence prevent viral infection. Viral infections that may be treated
 CC include human immunodeficiency virus (HIV), Epstein Barr virus (EBV),
 CC herpes simplex virus (HSV), human papilloma virus (HPV), adenovirus,
 CC respiratory syncytial virus (RSV), cytomegalovirus (CMV) and hepatitis
 CC B virus (HBV), especially HIV-1 infection.
 CC Sequence 36 BP: 0 A; 0 C; 24 G; 12 T;
 SQ

DE (TP)n internal repeat oligo used in synthetic HRGP gene modules.
 KW Synthetic gene; plant; gum; hydroxyproline-rich glycoprotein; HRGP;
 KW repetitive proline-rich protein; RRP; arabinogalactan protein; AGP;
 KW glycopeptide; ss.
 OS Synthetic.
 PN W09903978-A1.
 PD 28-JAN-1999.
 PE 21-JUL-1998; U15083.
 PR 20-JUL-1998; US-897556.
 PR 21-JUL-1997; US-897556.
 PA (UYOH-) UNIV OHIO.
 PI Kleiszewski MJ;
 PI WPI: 99-132225/11.
 PT Novel synthetic gene designed from repetitive peptide sequences - of
 PT hydroxyproline-rich glycoprotein
 PS Example 4; Page 38; 72pp; English.
 CC The invention relates to novel synthetic genes for plant gums. A new
 CC approach is described to the production of hydroxyproline-rich
 CC glycoproteins (HRGPs), repetitive proline-rich proteins (RPRs) and
 CC arabinogalactan proteins (AGPs). Synthetic genes comprising a nucleic
 CC acid encoding the peptide (Y01267) can be engineered for the production
 CC of repetitive glycopeptide modules in cells. The invention provided a new
 CC approach to the problem of producing plant gums that is not dependent on
 CC environmental factors and greatly simplifies the production of a variety
 CC of naturally occurring gums as well as designer gums.
 CC Sequence 42 BP: 13 A; 24 C; 1 G; 4 T;
 SQ

Query Match 76.7%; Score 13.8; DB 1; Length 42;
 Best Local Similarity 88.2%; Pred. No. 4.4e+02;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 ggtgtgtgtgtgtgtgt 17
 |||||
 DB 35 GGTGTGTGTGTGTGT 19
 RESULT 13
 V61397
 ID V61397 standard; DNA: 24 BP.
 AC V61397;
 DT 17-DEC-1998 (first entry)
 DE DNA template 5'-end region from W09844151 Example 1.
 KW Nucleic acid amplification; immobilised primer; sequencing;
 KW screening; diagnosis; in situ nucleic acid synthesis; fingerprinting;
 KW gene expression; ss.
 OS Synthetic.
 PN W09844151-A1.
 PD 08-OCT-1998.
 PE 01-APR-1998; G00961.
 PR 23-JUN-1997; GB-013238.
 PR 01-APR-1997; GB-006528.
 PR 01-APR-1997; GB-006529.
 PR 23-JUN-1997; GB-013236.
 PA (GLAXO) GLAXO GROUP LTD.
 PI Farinelli L, Kawashima E, Mayer P;
 PI WPI: 98-568282/48.
 PT New nucleic acid amplification - by extension of immobilised
 PT primers annealed to target, separation of strands, annealing
 PT extended primer to second primer and repeating extension
 PS Example 1; Page 23; 108pp; English.
 CC A method has been developed for nucleic acid (NA) amplification. The
 CC method comprises: (a) annealing single-stranded (ss) target NA to one of
 CC many primers (P), immobilised but having one end free for extension, and
 CC extending P using the target as template, producing an immobilised
 CC extension product (EP); (b) separating EP and target NA; (c) annealing
 CC EP to another P and extending this, using EP as template, to give a
 CC second extension product (EP2); and (d) optionally separating EP and EP2.
 CC The method is used to produce NA for sequencing, diagnosis and screening,
 CC as supports for other materials, for generating free NA (particularly in
 CC situ RNA synthesis), monitoring gene expression, identifying NA encoding
 CC rarely expressed gene products and identifying heterozygotes and in NA
 CC fingerprinting. A very high density of DNA-bearing areas can be randomly

CC arrayed and amplified on the support. The present sequence represents the
CC 5'-end of a DNA template used in an example from the present invention.
SQ Sequence 24 BP; 0 A; 0 C; 12 G; 12 T;

Query Match
Best Local Similarity 74.4%; Score 13.4; DB 1; Length 24;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 tctgtgtgtgtgctt 17
|||
Db 1 TGTGTGCTGTGCTT 15

RESULT 14
V07668/c
ID V07668 standard; DNA; 27 BP.
AC V07668:
DT 21-DEC-1998 (first entry)
DE Primer Mu-Bio used to amplify Mu-tagged sequence.
KW Bx1 gene; DIMBOA; 2,4-dihydroxy-7-methoxy-1,4 benzoxazin-3-one;
KW maize; benzoxazinone; insect resistance; herbicide resistance;
KW crop protection; transgenic plant; ss.
OS Synthetic.
PN W09840505-A1.
PD 17-SEP-1998.
PE 13-MAR-1998; 005078.
PR 13-MAR-1997; US-040513.
PA (DEKA-) DEKALB GENETICS CORP.
PI Chomet P, Frey M, Gierl A;
DR WPI; 98-520823/44.
PT New isolated maize Bx1 gene - used, e.g. for producing plants with
PT increased resistance to pests, pathogens or chemicals
PS Example 10; Page 94; 137pp; English.
CC Primers Mu-Bio, Msel sel/A(GCT), BfaI sel/A(GCT) and Mu Sel (see
CC V07668-71) were used to amplify Mu-tagged sequences in a Mu-tagged
CC Bx1 allele of maize. The Bx1 gene (see V07665) is involved in
CC the biosynthesis of the benzoxazinone DIMBOA. The gene provides a
CC tool for the production of transgenic plants with enhanced
CC expression profiles of benzoxazinone biosynthesis, and hence
CC resistance to pest infestation, herbicide damage and disease.
SQ Sequence 27 BP; 8 A; 10 C; 5 G; 3 T;

Query Match 74.4%; Score 13.4; DB 1; Length 27;
Best Local Similarity 82.4%; Pred. No. 6.4e+02;
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgctt 17
|||
Db 21 GGCGWTGGCGTGGCTT 5

RESULT 15
X32517/c
ID X32517 standard; DNA; 27 BP.
AC X32517:
DT 23-JUN-1999 (first entry)
DE Sequence of Mu-Bio.
KW Insertion; detection array; plant; screening; animal feed; bar gene;
KW pharmaceutical; industrial product; T-DNA; primer; ss.
OS Synthetic.
PN W09914373-A1.
PD 25-MAR-1999.
PE 17-SEP-1998; 019369.
PR 17-SEP-1997; US-932280.
PA (UTYA) UNITV YALE.
PI Dellaporta SL;
DR WPI; 99-229553/19.
PT Screening for insertion genomic mutations
PS Example 4; Page 43; 84pp; English.
CC The invention relates to detection of an insertion event in a genome by
CC screening at least one detection array comprising a DNA composition

CC enhanced for many insertion junctions. The method is used, particularly
CC for plants: (1) to determine the function of particular genes and (2) to
CC identify insertional mutations that confer a desirable property. The
CC genes or mutations identified may then be used to generate new plants for
CC uses as foods, animal feeds, e.g. as silage, or as sources of oil,
CC starch, pharmaceuticals, industrial products etc. Unlike known methods,
CC this process allows simultaneous screening of many unique insertional
CC events, and sequences of single members of a population can be identified
CC without individual analysis of each member of the population. The method
CC is generally applicable to animals, lower eukaryotes, prokaryotes and
CC particularly plants. Sequences X32516-522 represent primer and adapter
CC sequences. It is indicated that these sequences are in their -3'
SQ Sequence 27 BP; 8 A; 10 C; 5 G; 3 T;

Query Match 74.4%; Score 13.4; DB 1; Length 27;
Best Local Similarity 82.4%; Pred. No. 6.4e+02;
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgctt 17
|||
Db 21 GGCGWTGGCGTGGCTT 5

Search completed: June 4, 2000, 16:24:11
Job time: 28907 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:43 ; Search time 244.64 Seconds
(without alignments)
9.564 Million cell updates/sec

Title: US-09-164-714-13

Perfect score: 18
Sequence: 1 ggtgtgtgtgtgtgtttt 18

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/5E.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/5F.COMB.seq:*
7: /cgn2_6/ptodata/2/lna/5G.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.4	80.0	90	1	US-08-123-702-16 Sequence 16, App1
2	13.8	76.7	36	1	US-08-145-704-5 Sequence 5, App1
3	13.8	76.7	36	1	US-08-145-704-7 Sequence 7, App1
4	13.8	76.7	36	6	PCT-US93-07743-7 Sequence 7, App1
5	13.8	76.7	36	6	PCT-US93-07743-11 Sequence 11, App1
6	13.8	76.7	36	6	PCT-US93-07743-16 Sequence 16, App1
7	13.8	76.7	36	6	PCT-US96-11786-5 Sequence 5, App1
8	13.8	76.7	36	6	PCT-US96-11786-7 Sequence 7, App1
9	13.2	73.3	20	2	US-08-703-755A-4 Sequence 4, App1
10	13.2	73.3	24	1	US-08-145-704-16 Sequence 16, App1
11	13.2	73.3	24	6	PCT-US96-11786-16 Sequence 16, App1
12	13.2	73.3	39	2	US-08-411-727-11 Sequence 11, App1
13	13.2	73.3	45	1	US-08-197-770-6 Sequence 6, App1
14	13.2	73.3	45	1	US-08-197-770-8 Sequence 8, App1
15	13.2	73.3	46	4	US-08-495-695E-23 Sequence 23, App1
16	13.2	73.3	46	6	PCT-US94-14436-23 Sequence 23, App1
17	13.2	73.3	51	7	5474913-7 Patent No. 5474913
18	13.2	73.3	61	7	5474913-12 Patent No. 5474913
19	12.8	71.1	26	1	US-08-467-420A-43 Sequence 43, App1
20	12.8	71.1	26	2	US-08-470-110A-43 Sequence 43, App1
21	12.8	71.1	26	2	US-08-667-769A-43 Sequence 43, App1
22	12.8	71.1	26	6	US-08-940-371-43 Sequence 43, App1
23	12.8	71.1	26	6	PCT-US95-17082A-43 Sequence 43, App1
24	12.8	71.1	37	6	PCT-US93-07743-2 Sequence 2, App1
25	12.8	71.1	76	7	5182195-8 Patent No. 5182195
26	12.8	71.1	77	7	5182195-3 Patent No. 5182195
27	12.4	68.9	21	1	US-08-094-534-9 Sequence 9, App1

28	12.4	68.9	21	4	US-08-581-543-9 Sequence 9, App1
29	12.4	68.9	21	6	PCT-US94-08000-9 Sequence 9, App1
30	12.4	68.9	28	6	PCT-US92-09202-4 Sequence 4, App1
31	12.4	68.9	41	6	PCT-US92-09202-3 Sequence 3, App1
32	12.4	68.9	54	1	US-08-319-836B-1 Sequence 1, App1
33	12.4	68.9	54	6	PCT-US92-09202-2 Sequence 2, App1
34	12.4	68.9	54	6	PCT-US95-13142-1 Sequence 1, App1
35	12.4	68.9	62	1	US-08-319-836B-10 Sequence 10, App1
36	12.4	68.9	62	6	PCT-US95-13142-10 Sequence 10, App1
37	12.4	68.9	63	1	US-08-319-836B-11 Sequence 11, App1
38	12.4	68.9	63	6	PCT-US95-13142-11 Sequence 11, App1
39	12.4	68.9	85	1	US-08-276-452A-65 Sequence 65, App1
40	12.4	68.9	85	3	US-08-798-744-65 Sequence 65, App1
41	12.2	67.8	17	4	US-08-529-878B-35 Sequence 35, App1
42	12.2	67.8	26	1	US-08-149-096A-3 Sequence 3, App1
43	12.2	67.8	40	2	US-08-390-878-6 Sequence 6, App1
44	12.2	67.8	40	2	US-08-390-878-9 Sequence 9, App1
45	12.2	67.8	40	3	US-08-425-684-64 Sequence 64, App1

ALIGNMENTS

```
RESULT 1
US-08-123-702-16
; Sequence 16, Application US/08123702
; Patent No. 5604131
; GENERAL INFORMATION:
; APPLICANT: Wadsworth, Samuel
; APPLICANT: Snyder, Benjamin
; APPLICANT: Reddy, Verne, B.
; APPLICANT: Wei, Chamer
; TITLE OF INVENTION: A cDNA Genomic Hybrid Sequence Encoding App770
; Patent No. 5604131
; TITLE OF INVENTION: Containing a Genomic DNA Insert of the KI and OX-2 Regions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS: 45
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,702
; FILING DATE: 17-SEPT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TS1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-123-702-16

Query Match 80.0%; Score 14.4; DB 1; Length 90;
Best Local Similarity 93.8%; Pred. No. 1.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 3 tgggtgtgtgcttt 18
||||| |||||
DB 33 TCGTGTGTGCTTT 48

RESULT 2

US-08-145-704-5
; Sequence 5, Application US/08145704
; Patent No. 5567604
; GENERAL INFORMATION:
; APPLICANT: Rando, Robert F.
; APPLICANT: Rannewald, Susan
; APPLICANT: Zendegeul, Joseph G.
; APPLICANT: Joshua O. Ojwang
; TITLE OF INVENTION: Anti-Viral Guanosine-Rich
; TITLE OF INVENTION: Oligonucleotides
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.A.
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,704
; FILING DATE: 28-OCT-1993
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,027
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5574-CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/651-5151
; TELEFAX: 713/651-5246
; TELEX: 762829
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-145-704-5

Query Match 76.7%; Score 13.8; DB 1; Length 36;
Best Local Similarity 88.2%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgctt 17
||||| ||||| ||
DB 9 GGTGTGTGTGTGCTT 25

RESULT 3

US-08-145-704-7
; Sequence 7, Application US/08145704
; Patent No. 5567604
; GENERAL INFORMATION:
; APPLICANT: Rando, Robert F.
; APPLICANT: Rannewald, Susan
; APPLICANT: Zendegeul, Joseph G.
; APPLICANT: Joshua O. Ojwang

; TITLE OF INVENTION: Anti-Viral Guanosine-Rich
; TITLE OF INVENTION: Oligonucleotides
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.A.
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,704
; FILING DATE: 28-OCT-1993
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,027
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5574-CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/651-5151
; TELEFAX: 713/651-5246
; TELEX: 762829
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 36
; OTHER INFORMATION: /note="phosphorothioate backbone"

US-08-145-704-7

Query Match 76.7%; Score 13.8; DB 1; Length 36;
Best Local Similarity 88.2%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgctt 17
||||| ||||| ||
DB 9 GGTGTGTGTGTGCTT 25

RESULT 4

PCT-US93-07743-7
; Sequence 7, Application PC/TUS9307743
; GENERAL INFORMATION:
; APPLICANT: Jayaraman, Krishna
; APPLICANT: Vu, Huynh
; APPLICANT: Zendegeul, Joseph
; TITLE OF INVENTION: Cholesteryl-Modified Triple Helix
; TITLE OF INVENTION: Forming Oligonucleotides and uses thereof
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Adler/ Fulbright & Jaworski
; STREET: 1301 McKinney St Suite 4100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07743
FILING DATE: 19930817
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/934,065
FILING DATE: 21-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Adler Ph.D., Benjamin A.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5523
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-651-5587
TELEFAX: 713-651-5246
TELEX: 76-2829
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc_feature
LOCATION: 36
OTHER INFORMATION: /note="a NH2 group is attached to
OTHER INFORMATION: the 3' end"
PCT-US93-07743-7

Query Match 76.7%; Score 13.8; DB 6; Length 36;
Best Local Similarity 88.2%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgtt 17
||||| 11 11 11
DB 9 GGTGTGTGTGTGTGT 25

RESULT 5
PCT-US93-07743-11
Sequence 11, Application PC/TUS9307743
GENERAL INFORMATION:
APPLICANT: Jayaraman, Krishna
APPLICANT: Vu, Huynh
APPLICANT: Zendequal, Joseph
TITLE OF INVENTION: Cholesteryl-Modified Triple Helix
TITLE OF INVENTION: Forming Oligonucleotides and Uses Thereof
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. Adler/ Fulbright & Jaworski
STREET: 1301 McKinney St Suite 4100
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07743
FILING DATE: 19930817
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/934,065
FILING DATE: 21-AUG-1992
ATTORNEY/AGENT INFORMATION:

NAME: Adler Ph.D., Benjamin A.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5523
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-651-5587
TELEFAX: 713-651-5246
TELEX: 76-2829
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1
OTHER INFORMATION: /note="a cholesterol moiety is
OTHER INFORMATION: attached to the 3' end"
PCT-US93-07743-11

Query Match 76.7%; Score 13.8; DB 6; Length 36;
Best Local Similarity 88.2%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtt 17
||||| 11 11 11
DB 9 GGTGTGTGTGTGT 25

RESULT 6
PCT-US93-07743-16
Sequence 16, Application PC/TUS9307743
GENERAL INFORMATION:
APPLICANT: Jayaraman, Krishna
APPLICANT: Vu, Huynh
APPLICANT: Zendequal, Joseph
TITLE OF INVENTION: Cholesteryl-Modified Triple Helix
TITLE OF INVENTION: Forming Oligonucleotides and Uses Thereof
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. Adler/ Fulbright & Jaworski
STREET: 1301 McKinney St Suite 4100
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07743
FILING DATE: 19930817
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/934,065
FILING DATE: 21-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Adler Ph.D., Benjamin A.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5523
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-651-5587
TELEFAX: 713-651-5246
TELEX: 76-2829
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1
OTHER INFORMATION: /note="a propanolamine moiety is
OTHER INFORMATION: attached to the 3' end"
PCT-US93-07743-16

Query Match 76.7%; Score 13.8; DB 6; Length 36;
Best Local Similarity 88.2%; Pred. No.2.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtgtc 17
| | | | | | | | | | | | | | | | | | | | | |
Db 9 GGTGTGTGTGTGTGT 25

RESULT 7
PCT-US96-11786-5
; Sequence 5, Application PC/TUS9611786
; GENERAL INFORMATION:
; APPLICANT: Rando, Robert F.
; APPLICANT: Fennwald, Susan
; APPLICANT: Zendegui, Joseph G.
; APPLICANT: Ojwang, Joshua O.
; APPLICANT: Hogan, Michael E.
; APPLICANT: Pommer, Eyles
; APPLICANT: Mazunder, Abhijit
; TITLE OF INVENTION: Anti-viral Guanosine-Rich
; NUMBER OF SEQUENCES: 52
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/11786
; FILING DATE: 17-JULY-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/535,168; 60/001,505; 60/014,007; 60/013,688;
; APPLICATION NUMBER: 60/015,714; 60/016,271
; FILING DATE: 23-OCT-95; 17-JULY-96; 25-MARCH-96; 19-MARCH-96; 23-
; FILING DATE: APRIL-96; 17-APRIL-96
; ATTORNEY/AGENT INFORMATION:
; NAME: McDaniel, C. Steven
; REGISTRATION NUMBER: 33,962
; REFERENCE/DOCKET NUMBER: 1472-06214
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/238-8010
; TELEFAX: 713/238-8008
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US96-11786-5

Query Match 76.7%; Score 13.8; DB 6; Length 36;
Best Local Similarity 88.2%; Pred. No.2.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ggtgtgtgtgtgtc 17
| | | | | | | | | | | | | | | | | | | | | |
Db 9 GGTGTGTGTGTGTGT 25

RESULT 8
PCT-US96-11786-7
; Sequence 7, Application PC/TUS9611786
; GENERAL INFORMATION:
; APPLICANT: Rando, Robert F.
; APPLICANT: Fennwald, Susan
; APPLICANT: Zendegui, Joseph G.
; APPLICANT: Ojwang, Joshua O.
; APPLICANT: Hogan, Michael E.
; APPLICANT: Pommer, Eyles
; APPLICANT: Mazunder, Abhijit
; TITLE OF INVENTION: Anti-viral Guanosine-Rich
; NUMBER OF SEQUENCES: 52
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/11786
; FILING DATE: 17-JULY-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/535,168; 60/001,505; 60/014,007; 60/013,688;
; APPLICATION NUMBER: 60/015,714; 60/016,271
; FILING DATE: 23-OCT-95; 17-JULY-96; 25-MARCH-96; 19-MARCH-96; 23-
; FILING DATE: APRIL-96; 17-APRIL-96
; ATTORNEY/AGENT INFORMATION:
; NAME: McDaniel, C. Steven
; REGISTRATION NUMBER: 33,962
; REFERENCE/DOCKET NUMBER: 1472-06214
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/238-8010
; TELEFAX: 713/238-8008
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 36
; OTHER INFORMATION: /note="phosphorothioate
; OTHER INFORMATION: backbone"
; PCT-US96-11786-7

Query Match 76.7%; Score 13.8; DB 6; Length 36;
Best Local Similarity 88.2%; Pred. No.2.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 ggtgtgtgtgtgtc 17


```

; FILING DATE: APRIL-96, 17-APRIL-96
; ATTORNEY/AGENT INFORMATION:
; NAME: MCDANIEL, C. STEVEN
; REGISTRATION NUMBER: 33,962
; REFERENCE/DOCKET NUMBER: 1472-06214
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/238-8010
; TELEFAX: 713/238-8008
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US96-11786-16

Query Match      73.3%; Score 13.2; DB 6; Length 24;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgtgtt 18
Db 2 GGTGTGTGTGTGTGT 19

RESULT 12
US-08-411-727-11
; Sequence 11, Application US/08411727
; Patent No. 5705161
; Patent No. 5705161 5683703
; GENERAL INFORMATION:
; APPLICANT: VAN DER LEY, Peter Andre
; APPLICANT: POOLMAN, Jan Theunis
; APPLICANT: HOOGHERHOUT, Peter
; TITLE OF INVENTION: IMMUNOGENIC MENINGOCOCCAL LPS AND OTHER
; TITLE OF INVENTION: MEMBRANE VESICLES AND VACCINE THEREFROM
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street, Suite 200
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/411,727
; FILING DATE: 01-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL 9201716
; FILING DATE: 02-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/NL93/00163
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32925
; REFERENCE/DOCKET NUMBER: BO 38275
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-521-2297
; TELEFAX: 703-685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..34
; US-08-411-727-11

Query Match      73.3%; Score 13.2; DB 2; Length 39;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgtt 18
Db 11 GTTGTGTGTGTGTCTT 28

RESULT 13
US-08-197-770-6
; Sequence 6, Application US/08197770
; Patent No. 5514548
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Method for In Vivo Selection of
; TITLE OF INVENTION: Ligand-Binding Proteins
; NUMBER OF SEQUENCES: 23
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,770
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93 10 2484.8
; FILING DATE: 17-FEB-1993
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-197-770-6

Query Match      73.3%; Score 13.2; DB 1; Length 45;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgtt 18
Db 15 GGTGTGTGTGTGTATT 32

RESULT 14
US-08-197-770-8
; Sequence 8, Application US/08197770
; Patent No. 5514548
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Method for In Vivo Selection of
; TITLE OF INVENTION: Ligand-Binding Proteins
; NUMBER OF SEQUENCES: 23
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```


SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/197,770
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 93 10 2484.8
FILING DATE: 17-FEB-1993
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-197-770-8

Query Match 73.3%; Score 13.2; DB 1; Length 45;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgtgttt 18
||| ||||| ||| |||
Db 15 GGTGATGTGTGTGATTT 32

RESULT 15
US-08-495-695B-23
Sequence 23, Application US/08495695B
Patent No. 5976814
GENERAL INFORMATION:
APPLICANT: Bard, Jonathan A.
APPLICANT: Walker, Mary
APPLICANT: Brancheck, Theresa
APPLICANT: Weinsbank, Richard L.
TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE
TITLE OF INVENTION: Y/PEPTIDE YY/PANCREATIC POLYPEPTIDE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/495,695B
FILING DATE: 13-Jan-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44743-A-PCT-US/JFW/JHB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO

ANTI-SENSE: NO
US-08-495-695B-23

Query Match 73.3%; Score 13.2; DB 4; Length 46;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtgtgtgtgtgttt 18
||| ||||| ||| |||
Db 6 GGTGATGTGTGTGTCCTT 23

Search completed: June 4, 2000, 16:09:44
Job time: 28063 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:37 ; Search time 1236.38 Seconds

(without alignments)
-15.736 Million cell updates/sec

Title: US-09-164-714-14

Perfect score: 20
Sequence: 1 ccccttaaacatcgcac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenBank:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_om:*
5: gb_om:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pl2:*
11: gb_pl3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_om:*
23: em_om:*
24: em_om:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_ro:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_hg1:*
33: gb_hg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pl4:*
41: gb_hg3:*
42: gb_hg4:*
43: gb_hg5:*
44: gb_hg6:*

45: gb_hg7:*
46: em_hg1:*
47: em_hg2:*
48: em_hg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pl5:*
52: gb_hg8:*
53: gb_hg9:*
54: gb_hg10:*
55: gb_hg11:*
56: gb_hg12:*
57: gb_hg13:*
58: gb_hg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14.2	71.0	50	5	AR032933	AR032933 Sequence
2	14.2	71.0	50	5	I29673	I29673 Sequence 54
3	14.2	71.0	50	5	I91347	I91347 Sequence 54
4	13.8	69.0	69	5	A48319	A48319 Sequence 29
5	13.8	69.0	69	5	A48320	A48320 Sequence 30
6	13.6	68.0	55	1	TTPCT7KBA	K02338 T1 plasmid
7	13.2	66.0	76	10	HS030191	U30191 Human clone
8	12.6	63.0	22	5	AR066834	AR066834 Sequence
9	12.6	63.0	22	5	AR066998	AR066998 Sequence
10	12.6	63.0	33	5	A56675	A56675 Sequence 9
11	12.6	63.0	33	5	AR051609	AR051609 Sequence
12	12.2	61.0	46	5	AR032696	AR032696 Sequence
13	12.2	61.0	46	5	I29436	I29436 Sequence 30
14	12.2	61.0	46	5	I91110	I91110 Sequence 30
15	12.2	61.0	57	34	AF039609	AF039609 Drosophila
16	12.2	61.0	63	5	A18114	A18114 Oligonucleo
17	12.2	60.0	49	5	I56706	I56706 Sequence 22
18	12.2	60.0	50	5	AR032926	AR032926 Sequence
19	12.2	60.0	50	5	I29666	I29666 Sequence 53
20	12.2	60.0	50	5	I91340	I91340 Sequence 53
21	12.2	60.0	62	10	HC0146S12	U46970 Human type
22	12.2	60.0	72	9	HUMCOL4S11	D63526 Homo sapien
23	12.2	60.0	80	5	AR051980	AR051980 Sequence
24	12.2	60.0	80	5	I52154	I52154 Sequence 64
25	12.2	60.0	81	16	AF109762	AF109762 Hepatitis
26	12.2	60.0	94	13	HSPE33H11	AL009835 H. sapien
27	11.8	59.0	20	5	I81222	I81222 Sequence 6
28	11.8	59.0	23	5	AR055654	AR055654 Sequence
29	11.8	59.0	89	5	AR011127	AR011127 Sequence
30	11.6	58.0	20	5	AR066722	AR066722 Sequence
31	11.6	58.0	21	5	AR031094	AR031094 Sequence
32	11.6	58.0	24	5	A99109	A99109 Sequence 13
33	11.6	58.0	30	5	A14209	A14209 Oligonucleo
34	11.6	58.0	48	5	AR067776	AR067776 Sequence
35	11.6	58.0	48	5	E13195	E13195 Oligonucleo
36	11.6	58.0	59	5	E13419	E13419 Synthetic o
37	11.6	58.0	59	5	E13434	E13434 Synthetic o
38	11.6	58.0	60	5	AR006756	AR006756 Sequence
39	11.6	58.0	60	5	I71268	I71268 Sequence 6
40	11.6	58.0	78	5	I09608	I09608 Sequence 4
41	11.6	58.0	89	16	XCUT0893	U70893 Xestia c-ni
42	11.6	58.0	94	11	HUMIL75	M29052 Human inter
43	11.4	57.0	15	5	I77853	I77853 Sequence 56
44	11.4	57.0	23	24	E09203	E09203 Primer for
45	11.4	57.0	39	12	MUSBMP2AI	D45010 Mouse gene

ALIGNMENTS

RESULT 1
LOCUS AR032933 50 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 545 from patent US 5869241.
ACCESSION AR032933
VERSION AR032933.1 GI:5948538
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 50)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Method of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5869241-A 545 09-FEB-1999;
FEATURES
source Location/Qualifiers
1..50
BASE COUNT 13 a 7 c 16 g 14 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 5; Length 50;
Best Local Similarity 84.2%; Pred. No. 2.8e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgcca 19
||||| ||| ||| |||
DB 28 CCCCTTTAAATATCCCA 10

RESULT 2
LOCUS I29673 50 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 545 from patent US 5578444.
ACCESSION I29673
VERSION I29673.1 GI:1820464
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 50)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Sequence-directed DNA-binding molecules compositions and methods
JOURNAL Patent: US 5578444-A 545 26-NOV-1996;
FEATURES
source Location/Qualifiers
1..50
BASE COUNT 13 a 7 c 16 g 14 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 5; Length 50;
Best Local Similarity 84.2%; Pred. No. 2.8e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgcca 19
||||| ||| ||| |||
DB 28 CCCCTTTAAATATCCCA 10

RESULT 3
LOCUS I91347 50 bp DNA PAT 01-DEC-1998
DEFINITION Sequence 545 from patent US 5726014.
ACCESSION I91347
VERSION I91347.1 GI:3935817
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 50)

AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
TITLE Screening assay for the detection of DNA-binding molecules
JOURNAL Patent: US 5726014-A 545 10-MAR-1998;
FEATURES
source Location/Qualifiers
1..50
BASE COUNT 13 a 7 c 16 g 14 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 5; Length 50;
Best Local Similarity 84.2%; Pred. No. 2.8e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgcca 19
||||| ||| ||| |||
DB 28 CCCCTTTAAATATCCCA 10

RESULT 4
LOCUS A48319 69 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 29 from Patent WO9601900.
ACCESSION A48319
VERSION A48319.1 GI:2302112
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 69)
AUTHORS Layton,G.T., Garcia-Valcarcel,M., Fowler,W.J. and Harper,D.R.
TITLE IMMUNODOMINANT POLYPEPTIDES
JOURNAL Patent: WO 9601900-A 29 25-JAN-1996;
COMMENT BRITISH BIOTECH PHARM (GB)
Other publication AU 2803595 960209.
FEATURES
source Location/Qualifiers
1..69
BASE COUNT 23 a 9 c 18 g 19 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 5; Length 69;
Best Local Similarity 88.2%; Pred. No. 4.6e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ccccttaaacatcgcc 18
||||| ||| ||| |||
DB 64 CCCCTGAAACATCAC 48

RESULT 5
LOCUS A48320 69 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 30 from Patent WO9601900.
ACCESSION A48320
VERSION A48320.1 GI:2302113
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 69)
AUTHORS Layton,G.T., Garcia-Valcarcel,M., Fowler,W.J. and Harper,D.R.
TITLE IMMUNODOMINANT POLYPEPTIDES
JOURNAL Patent: WO 9601900-A 30 25-JAN-1996;
COMMENT BRITISH BIOTECH PHARM (GB)
Other publication AU 2803595 960209.
FEATURES
source Location/Qualifiers
1..69
BASE COUNT 19 a 18 c 9 g 23 t

ORIGIN

Query Match 69.0%; Score 13.8; DB 5; Length 69;
Best Local Similarity 88.2%; Pred. No. 4.6e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ccccttaaacatcgcc 18
||||| ||||| |||
DB 10 CCCTTGAAACATCACCC 26

RESULT 6
TIPTCTKBA 55 bp DNA BCT 13-MAR-1996
LOCUS TI Plasmid (from A.tumefaciens) T-region .7 kb transcript gene, 5'
DEFINITION flank.
ACCESSION K02338
VERSION K02338.1 GI:154732
KEYWORDS
SOURCE Plasmid TI (clone: pTIA6.) DNA.
ORGANISM plasmid TI
REFERENCE 1 (bases 1 to 55)
AUTHORS McPherson,J.C.
TITLE DNA sequence analysis of crown gall tumor T-DNA encoding the 0.7 kb transcript
JOURNAL Nucleic Acids Res. 12 (5), 2317-2325 (1984)
MEDLINE 84169535
FEATURES
Source location/Qualifiers
1..55
/organism="Plasmid TI"
/plasmid="Plasmid TI"
/specific_host="Agrobacterium tumefaciens"
/db_xref="taxon:2512"
/clone="pTIA6."

BASE COUNT 14 a 23 c 9 g 9 t

ORIGIN

Query Match 68.0%; Score 13.6; DB 1; Length 55;
Best Local Similarity 80.0%; Pred. No. 5.9e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgccac 20
||||| ||||| |||||
DB 7 CCCCCGGAATCATCGCCAC 26

RESULT 7
HSU30191 76 bp mRNA PRI 20-SEP-1996
LOCUS Human clone 1881-5 T-cell receptor delta chain mRNA, partial cds.
DEFINITION U30191
ACCESSION U30191.1 GI:940079
VERSION
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 76)
AUTHORS Maeruer,M.J., Martin,D., Walter,W., Liu,K., Zltvogel,L., Lotze,M.T., Haluszcak,K., Rabinowich,H., Dukesnoy,R., Storkus,W. and Human intestinal Vdelta+ lymphocytes recognize tumor cells of epithelial origin
J. Exp. Med. 183 (4), 1681-1696 (1996)
MEDLINE 96261669
REFERENCE 2 (bases 1 to 76)
AUTHORS Martin,D.M.
TITLE Direct Submision
JOURNAL Submitted (23-JUN-1995) Dina M. Martin, Surgery, Biochemistry, & Molecular Genetics, University of Pittsburgh, Desoto and O'Hara

FEATURES Streets, Pittsburgh, PA 15261, USA
source location/Qualifiers
1..76
/organism="Homo sapiens"
/isolate="patient #1881"
/db_xref="taxon:9606"
/clone="1881-5"
/sex="male"
/cell_type="cytotoxic T-lymphocyte"
/tissue_type="colon tumor"
/dev_stage="adult"
<1..>76
/codon_start=2
/product="T-cell receptor delta chain"
/protein_id="AAB08496.1"
/db_xref="GI:940080"
/translation="SWGTGQFLQIKGVGDPVKGQVDKL"

BASE COUNT 20 a 15 c 26 g 15 t

ORIGIN

Query Match 66.0%; Score 13.2; DB 10; Length 76;
Best Local Similarity 83.3%; Pred. No. 9.8e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgcc 18
||||| ||||| |||||
DB 58 CCCCTTACAGTATCGCC 41

RESULT 8
AR066834/c 22 bp DNA PAT 29-SEP-1999
LOCUS AR066834
DEFINITION Sequence 182 from patent US 5851760.
ACCESSION AR066834
VERSION AR066834.1 GI:5998056
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Evans,G.A. and Smith,M.W.
TITLE Method for generation of sequence sampled maps of complex genomes
JOURNAL Patent: US 5851760-A 182 22-DEC-1998;
FEATURES
source location/Qualifiers
1..22
/organism="unknown"

BASE COUNT 6 a 2 c 8 g 6 t

ORIGIN

Query Match 63.0%; Score 12.6; DB 5; Length 22;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgcca 19
||||| ||||| |||
DB 22 CCCCTTCAAAACAGTGCTA 4

RESULT 9
AR066998/c 22 bp DNA PAT 29-SEP-1999
LOCUS AR066998
DEFINITION Sequence 346 from patent US 5851760.
ACCESSION AR066998
VERSION AR066998.1 GI:5998220
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Evans,G.A. and Smith,M.W.
TITLE Method for generation of sequence sampled maps of complex genomes

JOURNAL Patent: US 5851760-A 346 22-DEC-1998;
FEATURES Location/Qualifiers
source 1..22
/organism="unknown"
BASE COUNT 4 a 4 c 7 g 7 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 5; Length 22;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ccccttaaacatgccac 20
1 | | | | | | | | | |
Db 21 CTCCTTAGAACATAGCCAC 3

RESULT 10
A56675 33 bp DNA PAT 03-MAR-1998
LOCUS
DEFINITION Sequence 9 from Patent EP0733704.
ACCESSION A56675
VERSION A56675.1 GI:3712718
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 33)
AUTHORS Boeck, A.P., Mayer, D., Schlensoy, V.D. and Candussio, A.D.
TITLE Controllable expression system
JOURNAL Patent: EP 0733704-A 9 25-SEP-1996;
CONSORTIUM ELEKTROCHEM IND (DE)
COMMENT Other publication FI 961323 960925
Other publication JP 8266291 961015
Other publication DE 19514056 961017
Other publication CA 2171668 960925.
Location/Qualifiers

FEATURES
source 1..33
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 11 a 11 c 7 g 4 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 5; Length 33;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ccccttaaacatgccca 19
1 | | | | | | | | | |
Db 7 CCCGCTTAACATCTCCA 25

RESULT 11

AR051609 33 bp DNA PAT 29-SEP-1999
LOCUS
DEFINITION Sequence 9 from patent US 5830692.
ACCESSION AR051609
VERSION AR051609.1 GI:5974973
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 33)
AUTHORS Boock, A., Mayer, D., Schlensoy, V. and Candussio, A.
TITLE Expression system which can be regulated
JOURNAL Patent: US 5830692-A 9 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..33
/organism="unknown"

BASE COUNT 11 a 11 c 7 g 4 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 5; Length 33;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ccccttaaacatgccca 19
1 | | | | | | | | | |
Db 7 CCCGCTTAACATCTCCA 25

RESULT 12
AR032696 46 bp DNA PAT 29-SEP-1999
LOCUS
DEFINITION Sequence 308 from patent US 5869241.
ACCESSION AR032696
VERSION AR032696.1 GI:5948301
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 46)
AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M., Turin, L.M. and Fry, K.E.
TITLE Method of determining DNA sequence preference of a DNA-binding
JOURNAL molecule
Patent: US 5869241-A 308 09-FEB-1999;
FEATURES Location/Qualifiers
source 1..46
/organism="unknown"

BASE COUNT 8 a 19 c 12 g 7 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 5; Length 46;
Best Local Similarity 82.4%; Pred. No. 3.4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatgcc 17
1 | | | | | | | | | |
Db 17 CCCCTTAAGCGCGC 33

RESULT 13
I29436 46 bp DNA PAT 06-FEB-1997
LOCUS
DEFINITION Sequence 308 from patent US 5578444.
ACCESSION I29436
VERSION I29436.1 GI:1820227
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 46)
AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M., Turin, L.M. and Fry, K.E.
TITLE Sequence-directed DNA-binding molecules compositions and methods
JOURNAL Patent: US 5578444-A 308 26-NOV-1996;
FEATURES Location/Qualifiers
source 1..46
/organism="unknown"

BASE COUNT 8 a 19 c 12 g 7 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 5; Length 46;
Best Local Similarity 82.4%; Pred. No. 3.4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatgcc 17
1 | | | | | | | | | |
Db 17 CCCCTTAAGCGCGC 33

RESULT 14
191110

```

LOCUS       191110               46 bp    DNA
DEFINITION  Sequence 308 from patent US 5726014.
ACCESSION   191110
VERSION     191110.1  GI:3935580
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 46)
AUTHORS    Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turlin,L.M.
TITLE      Screening assay for the detection of DNA-binding molecules
JOURNAL    Patent: US 5726014-A 308 10-MAR-1998;
FEATURES
             Location/Qualifiers
             source             1..46
                                 /organism="unknown"
BASE COUNT   8 a          19 c          12 g          7 t
ORIGIN
Query Match      61.0%; Score 12.2; DB 5; Length 46;
Best Local Similarity 82.4%; Pred. No. 3.4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      1  ccccttaaacatgc 17
        |||||
Db      17  CCCCTTTAAAGCGCGC 33

RESULT  15
AF039609  57 bp    DNA          INV          23-MAR-1998
LOCUS     AF039609/c
DEFINITION Drosophila capricorni xanthine dehydrogenase gene, Intron II.
ACCESSION AF039609
VERSION   AF039609.1  GI:2981324
KEYWORDS
SOURCE    Drosophila capricorni.
ORGANISM  Drosophila capricorni.
           Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
           Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
           Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 57)
AUTHORS   Tarrío,R., Rodríguez-Trelles,F. and Ayala,F.J.
TITLE     New drosophila introns originate by duplication
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (4), 1658-1662 (1998)
MEDLINE   98132647
REFERENCE 2 (bases 1 to 57)
AUTHORS   Tarrío,R., Rodríguez-Trelles,F. and Ayala,F.J.
TITLE     Direct Submission
JOURNAL   Submitted (22-DEC-1997) Genetica, Universitat Autònoma de
           Barcelona, Barcelona, CA 08193, Spain
FEATURES
             Location/Qualifiers
             source             1..57
                                 /organism="Drosophila capricorni"
                                 /db_xref="taxon:46840"
                                 1..57
                                 /note="Intron II; xanthine dehydrogenase"
BASE COUNT   16 a          16 c          12 g          21 t
ORIGIN
Intron
Query Match      61.0%; Score 12.2; DB 34; Length 57;
Best Local Similarity 82.4%; Pred. No. 3.4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      4  cttaaacatgcgcac 20
        | ||||| |||||
Db     21  CATTAAGATTCGCGAC 5

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:11 ; Search time 322.35 Seconds
(without alignments)
15.523 Million cell updates/sec

Title: US-09-164-714-14

Perfect score: 1 ccccttaaacatcgccac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 08
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.2	71.0	50	1 Q69795	Epstein Barr Virus
2	14.2	71.0	50	1 T64257	EBV R2 (start site
3	14.2	71.0	50	1 X17545	Test sequence from
4	13.8	69.0	69	1 T11012	Coding sequence fo
5	13.6	69.0	42	1 X33064	Human BRCA2 gene P
6	13.4	67.0	36	1 X04254	MMP-7 linker regio
7	13	65.0	21	1 X041535	Antisense oligomer
8	12.8	64.0	50	1 X15997	PCR primer C1-7 us
9	12.8	64.0	50	1 X16000	PCR primer C2-2 us
10	12.8	64.0	50	1 Q90839	Wild type hMLH1 ge
11	12.6	63.0	22	1 Q82182	Chromosome 11 (loc
12	12.6	63.0	22	1 Q82346	Chromosome 11 (loc
13	12.6	63.0	33	1 T33684	Primer used to amp
14	12.4	62.0	60	1 V09404	Human ICE LAP-6 cd
15	12.2	61.0	28	1 X02116	Human FEN-1 PCR pr
16	12.2	61.0	46	1 Q69558	Human thrombospond
17	12.2	61.0	46	1 T64020	Human thrombospond
18	12.2	61.0	46	1 X17308	Test sequence from
19	12	60.0	49	1 Q99620	Human Ah receptor
20	12	60.0	50	1 Q69788	Epstein Barr Virus
21	12	60.0	50	1 T64250	EBV L1 1 kb gene (
22	12	60.0	50	1 X17538	Test sequence from
23	12	60.0	66	1 V09407	Human ICE LAP-6 cd
24	12	60.0	80	1 Q92534	Forward primer AFl
25	11.8	59.0	20	1 T95216	PCR primer 1 for t
26	11.8	59.0	25	1 V37959	ECCEPO section 1 co
27	11.8	59.0	25	1 V38003	SCEPO section 1 co
28	11.8	59.0	75	1 V75933	Staphylococcus aur
29	11.6	58.0	20	1 Q82070	Chromosome 11 (loc
30	11.6	58.0	21	1 Q87800	Primer pair 10B D3
31	11.6	58.0	24	1 X27972	PCR primer for B18
32	11.6	58.0	34	1 Q75834	Sense primer to am
33	11.6	58.0	48	1 V02491	Bridge oligomer MP
34	11.6	58.0	48	1 V02470	Primer MP92 for a

35	11.6	58.0	51	1 Q87671	Human superoxide d
36	11.6	58.0	51	1 V73847	Human SOD1 derived
37	11.6	58.0	58	1 T94059	Oligo used in cons
38	11.6	58.0	58	1 T94061	Oligo used in cons
39	11.6	58.0	60	1 T95110	Monkey Immunoglobu
40	11.6	58.0	74	1 T63629	Probe for L-prolin
41	11.6	58.0	78	1 Q03124	DNA sequence of th
42	11.4	57.0	15	1 T57364	RSV N hammerhead r
43	11.4	57.0	23	1 Q74017	Human Interleukin
44	11.4	57.0	25	1 V42672	BCSG1 gene 5' prim
45	11.4	57.0	31	1 X38898	Human genomic DNA

ALIGNMENTS

RESULT 1	
ID 069795/c	
AC 069795;	
DE 06-MAR-1995 (first entry)	
KW Epstein Barr Virus R2 (start site 88897), target region.	
KM DNA protein-binding assay; test sequence; screening sequence;	
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;	
KM origin of replication; UL9; transcription factor; TFIID; ds.	
OS Synthetic.	
PN WO9414980-A.	
PD 07-JUL-1994.	
PF 20-DEC-1993: U12388.	
PR 23-DEC-1992: US-996783.	
PR 17-SRP-1993: US-123936.	
PA (GENE-) GENELABS TECHNOLOGIES INC.	
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;	
DR WPI: 94-234711/28.	
PT Sequence-directed DNA-binding molecules - useful in	
PS pharmaceuticals and as molecular reagents	
PS Claim 28, page 484; 587pp; English.	
CC A DNA protein-binding assay is provided, useful for screening	
CC libraries of synthetic or biological cpds. for their ability	
CC to bind DNA test sequences. The assay is versatile in that any	
CC number of test sequences can be tested by placing the test sequence	
CC adjacent to a defined protein-binding screening sequence. Binding	
CC of mols. to these test sequences changes the binding characteristics	
CC of the protein mol. to its cognate binding sequence. When such a mol.	
CC binds the test sequence, the equilibrium of the DNA:protein complexes	
CC is disturbed, generating changes in the concentration of free DNA probe.	
CC One application of this method is to eucaryotic general transcription	
CC factors (e.g. TFIID), where the target region is typically selected	
CC from DNA sequences adjacent to the binding site for the eucaryotic	
CC transcription factor. Numerous exemplary test sequences are given:	
CC the sequences in Q69251-731 and Q69850 correspond to promoter targets	
CC (typically, TATA box-contg. sites) for human genes and the sequences in	
CC Q69732-849 correspond to promoter targets for viral genes. The test	
CC sequences may also be randomly generated. DNA:protein interaction may	
CC be used for screening purposes, e.g. the Herpes Simplex Virus (HSV)	
CC origin of replication and UL9 (see Q69851-52, Q69865 and Q69891).	
SQ Sequence 50 BP; 13 A; 16 G; 14 T;	
Query Match 71.0%; Score 14.2; DB 1; Length 50;	
Best Local Similarity 84.2%; Pred. No. 88;	
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
OY 1 ccccttaaacatcgcca 19	
DB 28 CCCCTTTTAATATCCCA 10	
RESULT 2	
ID T64257/c	
AC T64257;	
DR 17-MAR-1997 (first entry)	

DE EBV R2 (start site 88897) TFIID binding site.
 KM Duplex DNA; target region; binding characteristic; DNA binding protein;
 KM TFIID; transcription factor; binding site; inhibition; enhance;
 KM cancer; inherited genetic disorder; ds.
 OS Epstein-Barr virus.
 PN US578444-A.
 PD 26-NOV-1996.
 PF 27-JUN-1991; 723618.
 PR 23-DEC-1992; US-723618.
 PR 17-SEP-1993; US-123936.
 PR 20-DEC-1993; US-171389.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
 DR WPI: 97-020402/02.
 PT Altering binding characteristics of DNA binding proteins to duplex
 PT DNA - by attaching specific small cpd. to target region close to the
 PT protein's binding site, useful in treatment of viral disease, cancer
 PT etc.
 PS Claim 6; Column 377; 264pp; English.
 CC The sequences given in T63713-4312 represent duplex DNA's which act
 CC as target regions in the method of the invention. The method for
 CC altering the binding characteristics of a DNA-binding protein to duplex
 CC DNA comprises contacting the duplex DNA with a small molecule which
 CC binds sequence-specifically to a target region, where, when the small
 CC molecule is bound to the target region, it is adjacent to, but not
 CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
 CC The small molecule is added at a concentration effective to alter the
 CC binding of the DNA binding protein, pref. TFIID, to its binding site on
 CC the duplex DNA. The binding of the small molecule may inhibit or enhance
 CC the binding of the DNA-binding protein to its binding site. The
 CC compounds isolated using this method are potentially useful as
 CC therapeutic agents for treatment of any disease which involves a
 CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
 CC The method is suitable for screening large biological or chemical
 CC libraries and allows determination of sequence-specific and relative
 CC affinities of known DNA-binding agents for different DNA sequences.
 CC The design of these duplex DNA's allows a single DNA:protein interaction
 CC to be used for screening sequence-specific, or preferential, DNA binding
 CC proteins that recognise almost any possible sequence (see also T49539-
 CC 74).
 CC Sequence 50 BP; 13 A; 7 C; 16 G; 14 T;
 SQ

Query Match 71.0%; Score 14.2; DB 1; Length 50;
 Best Local Similarity 84.2%; Pred. No. 88;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatgcga 19
 ||||| ||| ||| |||
 Db 28 CCCCTTTAAATATCCCA 10

RESULT 3
 X17545/c
 AC X17545;
 DT 06-MAY-1999 (first entry)
 DE Test sequence; DNA-binding molecule; screening sequence; human;
 KM nucleic acid amplification; target; viral; ds.
 OS Epstein-Barr virus.
 PN US5869241-A.
 PD 09-FEB-1999.
 PF 07-JUN-1995; 475228.
 PR 20-DEC-1993; US-171389.
 PR 27-JUN-1991; US-723618.
 PR 23-DEC-1992; US-996783.
 PR 17-SEP-1993; US-123936.
 PR 07-JUN-1995; US-475228.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
 DR WPI: 99-152755/13.

PT Determination of DNA sequence preference of a DNA-binding molecule -
 PT based on inhibition of binding of protein to oligonucleotide
 PT sequence attached to test sequence
 PS Claim 3; Columns 379-380; 270pp; English.
 CC Sequences X17001 to X17600 represent specifically claimed target test
 CC sequences that are used in the method of the invention of determining the
 CC DNA sequence preference of a DNA-binding molecule. The method comprises:
 CC (i) adding a test molecule and a DNA-binding protein to a mixture of
 CC duplex DNA test oligonucleotides, each of the test oligonucleotides
 CC having a test sequence adjacent to a screening sequence, where the
 CC screening sequence binds to the DNA-binding protein with a binding
 CC affinity that is independent of the DNA sequence of the test sequence,
 CC and where the mixture of duplex DNA test oligonucleotides includes
 CC several test sequences; (ii) incubating the test molecule, the mixture of
 CC duplex DNA test oligonucleotides and the DNA-binding protein for a time
 CC sufficient to permit binding of the test molecule to test sequences in
 CC the duplex DNA; (iii) separating unbound test oligonucleotides from test
 CC oligonucleotides bound to binding protein; (iv) amplifying the unbound
 CC test oligonucleotides; (v) repeating steps (ii) to (iv); (vi) isolating
 CC the amplified test oligonucleotides; and (vii) sequencing the isolated
 CC test oligonucleotides. Test sequences X17001-X17481 and X17600 correspond
 CC to promoter targets for human genes and test sequences X17482-X17599
 CC correspond to promoter targets for viral genes.
 CC Sequence 50 BP; 13 A; 7 C; 16 G; 14 T;
 SQ

Query Match 71.0%; Score 14.2; DB 1; Length 50;
 Best Local Similarity 84.2%; Pred. No. 88;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatgcga 19
 ||||| ||| ||| |||
 Db 28 CCCCTTTAAATATCCCA 10

RESULT 4
 T11012/c
 ID T11012 standard; DNA: 69 BP.
 AC T11012;
 DT 09-SEP-1996 (first entry)
 DE Coding sequence for Varicella zoster gE glycoprotein residues 141-161.
 KM Glycoprotein; gE; VZV; antibody; gB; Immunoreactive; Immune response;
 KM infection; diagnosis; therapy; ds.
 OS Varicella-zoster virus.
 PI Key location/Qualifiers
 FH 1..4
 FT misc-feature /*tag= a
 FT /*note= "5' overhang"
 FT /*tag= b
 FT /*note= "5' overhang GATC"
 FT misc-feature
 FT
 FT
 FT
 PN WO9601900-A1.
 PD 25-JAN-1996.
 PF 03-JUL-1995; G01566.
 PR 07-JUL-1994; GB-013751.
 PA (BRBI-) BRITISH BIOTECH PHARM LTD.
 PI Fowler MJ, Garcia-valcarcel Munoz-repiso M, Harper DR;
 PI Layton GT;
 DR WPI: 96-097630/10.
 DR P-ESDB: R88536.
 PT New isolated Varicella zoster gE polypeptide(s) - used to develop
 PT products for use in vaccines, passive immunisation and diagnosis
 PT involving VZV infection
 PS Example 6; Page 37; 47pp; English.
 CC This sequence represents the coding sequence for residues 141-161 of the
 CC Varicella-zoster virus (VZV) gE glycoprotein. The encoded sequence (and
 CC the sequences represented by R88522-R88535 and R88537-R88549) are used to
 CC create antibodies against the VZV gE glycoprotein. gE is one of six
 CC glycoproteins encoded by the VZV genome. From these six proteins, gE and
 CC gB are the major immunoreactive glycoproteins. These sequences can be
 CC used for stimulating an immune response against VZV infection. These
 CC peptides can also be used for determining the presence of anti-VZV gE
 CC antibodies in a sample, and in the diagnosis of VZV infection. The

CC antibodies against these sequences can be used for passive immunisation
CC treatment, and in diagnostic applications. This sequence contains the
CC major VZV GE immunodominant epitope and allows the development of
CC products which can produce an enhanced and broader immune response.
SQ Sequence 69 BP; 23 A; 9 C; 18 G; 19 T;

Query Match 69.0%; Score 13.8; DB 1; Length 69;
Best Local Similarity 88.2%; Pred. No. 1.5e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cccttaaacatcgcc 18
||||| ||||||| ||
Db 64 CCCTTGAACACATCACC 48

RESULT 5
X33064/C
ID X33064 standard; DNA: 42 BP.
AC X33064:
DT 21-JUN-1999 (first entry)
DE Human BRCA2 gene PCR primer SEQ ID NO:77.
KW Human; BRCA2; genetic testing; protein therapy; haplotype; detection;
KW gene therapy; breast cancer; ovarian cancer; PCR primer; ss.
OS Synthetic.
ON Homo sapiens.
PN MO9090164-A1.
PD 25-FEB-1999.
PE 14-AUG-1998; 016905.
PR 22-MAY-1998; US-084471.
PR 15-AUG-1997; US-055784.
PR 07-NOV-1997; US-064926.
PR 12-NOV-1997; US-065367.
PR 01-MAY-1998; US-071715.
PA (ONCO-) ONCOMED INC.
PI Eskandari T, Jackson GM, Murphy PD, Olson SJ, Park M,
PI Rabin MB, Schryer B, White MB, Yoshikawa M;
DR WPI: 99-190163/16.
PT New coding sequence haplotypes of the human BRCA2 gene - used to
PT develop products for determining susceptibility to, detection and
PT treatment of breast or ovarian cancer
PS Example 1; Page 34; 226pp; English.
CC The present invention describes genomic DNA which contains a BRCA2 gene
CC where the first 12 nucleotides beginning exon 5 are 5'-TCCGTGTTCT-3'
CC as in sequence (I) (see X03249), where nucleotides numbers
CC 5782-5790 are GTTGTGTT as in sequence (IV) (see X30253), and where the
CC last 20 nucleotides encoding exon 15 are 5'-CTGCGTGTCTCAATAACAG-3' as
CC in sequence (II) (see X30251) and the first 20 nucleotides beginning
CC exon 16 are 5'-CTGTATACGTATGCGCTTTC-3' as in sequence (III) (see X30253).
CC Products and methods from the present invention can be used for
CC identifying mutations in the BRCA2 gene leading to predisposition or
CC higher susceptibility to breast or ovarian cancer. They can also be used
CC for detection and gene therapy for breast and ovarian cancers. They can
CC be used in methods for monitoring disease progression, for determining
CC patients suited for gene and protein replacement progression, or for
CC detecting the presence or quantifying the amount of a tumour growth
CC inhibitor following such therapy. The BRCA2 protein, polypeptides, their
CC functional equivalents, antibodies, and PMS may also be useful in the
CC study of the characteristics of BRCA2 proteins, such as structure and
CC function of BRCA2 in oncogenesis or subcellular localisation of BRCA2
CC protein in normal and cancerous cells. X33001 to X33097 represent
CC PCR primers used in the amplification of the human BRCA2 gene.
SQ Sequence 42 BP; 14 A; 5 C; 12 G; 11 T;

Query Match 68.0%; Score 13.6; DB 1; Length 42;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 cccttaaacatcgccac 20
||| ||||||| ||| ||
Db 36 CCACTTAAACACCACTAC 17

RESULT 6

ID X04254/C
ID X04254 standard; DNA: 36 BP.
AC X04254:

DT 15-APR-1999 (first entry)
DE MWP-7 linker regions of PAP-217.
KW Ricin-like toxin; cancer; viral infection; parasitic infection;
KW linker; B chain; A chain; protease; fungal infection; malaria;
KW leucocyte proliferation; cytomegalovirus; herpes; hepatitis;
KW rhinovirus; laryngotracheitis; poliomyelitis; varicella zoster;
KW cystic fibrosis; multiple sclerosis; ds.
OS unidentified.
ON Synthetic.
PN MO9849311-A2.
PD 05-NOV-1998.
PE 30-APR-1998; CA0394.
PR 29-OCT-1997; US-063715.
PR 30-APR-1997; US-045148.
PA (DNOV-) DE NOVO ENZYME CORP.

PI Borgford T;
DR WPI: 99-009431/01.
PT New nucleic acid encoding ricin-like toxin with an interchain linker
PT cleaved by protease - is specific for diseased cells, useful for,
PT e.g. killing selectively cancer or infected cells
PS Disclosure, Fig 4b; 352pp; English.
CC The present invention describes new purified and isolated nucleic acids
CC (I) encoding: (1) the A and B chains of a ricin-like toxin (II); and
CC (11) a heterologous linker, joining the two chains and including a
CC cleavage recognition site for a disease-specific protease (III). Also
CC described are: (1) plasmids or baculovirus transfer vectors that contain
CC (I); and (2) recombinant protein (IV) consisting of the A and B chains
CC of (II) joined by the specified linker. (IV), produced by expression of
CC (I) in host cells, are used to inhibit or kill diseased cells that
CC produce (III), particularly for treating cancers (e.g. leucocyte
CC proliferation; cancer of ovary, pancreas, breast or prostate; glioma) or
CC infections caused by fungi, parasites (e.g. malaria) or viruses (e.g.
CC cytomegalovirus (CMV), herpes, hepatitis, rhinovirus, laryngotracheitis,
CC poliomyelitis or varicella zoster), also cystic fibrosis and multiple
CC sclerosis. Alternatively, (I) is used to express (IV) in vivo. (IV) is
CC toxic specifically for (III)-expressing cells and does not depend for
CC specificity on a cell-binding component. When used to treat virus-
CC infected cells, transcytosis and cytotoxicity of (IV) are increased by
CC retrograde translocation from endoplasmic reticulum to cytoplasm (which
CC some viruses exploit to avoid immune detection), so selectivity and
CC safety are further improved. (IV) are not toxic until chain A is
CC released and this occurs only in target cells. The present sequence
CC represents a nucleotide sequence from the present invention.
SQ Sequence 36 BP; 6 A; 7 C; 9 G; 14 T;

Query Match 67.0%; Score 13.4; DB 1; Length 36;
Best Local Similarity 93.3%; Pred. No. 2.2e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 ttaaacatcgccac 20
||||| ||||||| ||
Db 35 TTAAACTGCGCAC 21

RESULT 7

Q41535
ID Q41535 standard; DNA: 21 BP.
AC Q41535:

DT 10-AUG-1993 (first entry)
DE Antisense oligomer targeting BHRF2 (Anpase/ori binding) AUG site.
KW Epstein Barr virus; EBV; hybridisation; antisense modulator;
KW replication; nasopharyngeal carcinoma; Burkitt lymphoma; Sjogren's;
KW syndrome; infectious mononucleosis; latent; active; infection; ss.
ON Synthetic.
PN MO9307882-A.
PD 29-APR-1993.
PF 23-OCT-1992; U08989.

PR 25-OCT-1991: US-783605.
 PA (ISIS-) ISIS PHARM INC.
 PI Anderson KP, Eckert DJ;
 DR WPI: 93-152174/18.
 PT Oligo:nucleotide(s) hybridising with RNA of Epstein Barr virus -
 for treating active, latent and chronic EBV infections and
 associated diseases e.g. nasopharyngeal carcinoma, Burkitt's
 lymphoma
 PS Claim 1: Page 21: 45pp; English.
 CC The synthetic peptide is an antisense modulator of Epstein Barr
 virus and pref. contains at least one phosphorothioate linking gp.
 CC and modifications in the 2' position. These modifications improve
 CC penetration into regions of the cell that contain nucleic acid and
 CC also resistance to degradation by nucleases. The oligonucleotide
 CC targets the AUG site of BBLF2 (Arapase/ori binding) and hybridises thus
 CC inhibiting replication of EBV. The oligomer may be used for
 CC treating or preventing EBV-associated diseases, e.g. nasopharyngeal
 CC carcinoma, Burkitt's lymphoma, Sjogren's syndrome, infectious
 CC mononucleosis etc. The oligomer is effective against both latent and
 CC active EBV infection.
 CC See also Q40575-9 and Q41517-44.
 SQ Sequence 21 BP; 6 A; 7 C; 3 G; 5 T;

Query Match 65.0%; Score 13; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ccttaaacatc 15
 |||||
 DB 1 CCTTTAAACATC 13

RESULT 8
 ID X15997/c
 AC X15997: standard; DNA: 50 BP.
 DT 14-MAY-1999 (first entry)
 DE PCR primer C1-7 used for construction of a synthetic FAOD-P cDNA.
 KW Fructosylamino acid oxidase; FAOD-P; Penicillium janthinellum S-3413;
 KW PCR primer: ss.
 OS Synthetic.
 OS Penicillium janthinellum.
 PN J11046769-A.
 PD 23-FEB-1999.
 PE 05-AUG-1997: 210609.
 PR 05-AUG-1997: JP-210609.
 PA (KYOT-) KYOTO DAICHI KAGAKU KK.
 DR WPI: 99-208112/18.
 PT New DNA coding fructosylamino acid oxidase - synthetically designed
 PT and recombinantly prepared
 PS Example 1: Fig 4; 21pp; Japanese.
 CC PCR primers X15951-6006 were used to construct a synthetic cDNA
 CC sequence encoding a fructosylamino acid oxidase (FAOD-P). The
 CC wild type FAOD-P sequence is derived from Penicillium
 CC janthinellum S-3413.
 SQ Sequence 50 BP; 15 A; 8 C; 14 G; 13 T;

Query Match 64.0%; Score 12.8; DB 1; Length 50;
 Best Local Similarity 87.5%; Pred. No. 4.6e+02;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 tttaaacatgcac 20
 |||||
 DB 32 TTTTAAAGCATCACAC 17

RESULT 9
 ID X16000
 AC X16000: standard; DNA: 50 BP.
 DT 14-MAY-1999 (first entry)

DE PCR primer C2-2 used for construction of a synthetic FAOD-P cDNA.
 KW Fructosylamino acid oxidase; FAOD-P; Penicillium janthinellum S-3413;
 KW PCR primer: ss.
 OS Synthetic.
 OS Penicillium janthinellum.
 PN J11046769-A.
 PD 23-FEB-1999.
 PE 05-AUG-1997: 210609.
 PR 05-AUG-1997: JP-210609.
 PA (KYOT-) KYOTO DAICHI KAGAKU KK.
 DR WPI: 99-208112/18.
 PT New DNA coding fructosylamino acid oxidase - synthetically designed
 PT and recombinantly prepared
 PS Example 1: Fig 4; 21pp; Japanese.
 CC PCR primers X15951-6006 were used to construct a synthetic cDNA
 CC sequence encoding a fructosylamino acid oxidase (FAOD-P). The
 CC wild type FAOD-P sequence is derived from Penicillium
 CC janthinellum S-3413.
 SQ Sequence 50 BP; 12 A; 12 C; 9 G; 17 T;

Query Match 64.0%; Score 12.8; DB 1; Length 50;
 Best Local Similarity 87.5%; Pred. No. 4.6e+02;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 tttaaacatgcac 20
 |||||
 DB 27 TTTTAAAGCATCACAC 42

RESULT 10
 ID Q90839
 AC Q90839: standard; cDNA: 92 BP.
 DT 04-MAR-1996 (first entry)
 DE Wild type hMLH1 gene exon 6.
 KW hMLH1; wild type; MutL homologue; cancer diagnosis; mismatch repair;
 KW tumour; susceptibility; mutation detection; exon 6; ss.
 OS Homo sapiens.
 PN W09516793-A1.
 PD 22-JUN-1995.
 PE 16-DEC-1994: U14746.
 PR 17-DEC-1995: US-168877.
 PR 08-MAR-1994: US-209521.
 PR 09-DEC-1994: US-352902.
 PA (DAND) DANA FARBER CANCER INST INC.
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 PI Baker SM, Bollaag RJ, Bronner CE, Kolodner RD, Liskay RM;
 DR WPI: 95-231583/30.
 PT Determn. of a mutation in a mutL homologue or gene prod. in a tissue
 PT - used to diagnose cancer susceptibility, and to identify and
 PT classify a DNA mismatch-repair-defective tumour
 PS Claim 10: Fig 4A-2; 16pp; English.
 CC Q90839 is the wild type hMLH1 (a MutL homologue) gene exon 6. A
 CC mutation in an analogous segment of a hMLH1 or hPMS1 nucleic acid
 CC isolated from a subject, can be detected by comparing it with the
 CC above gene fragment. This method can be used to diagnose cancer
 CC susceptibility, or to identify and classify a DNA mismatch-repair
 CC defective tumour.
 SQ Sequence 92 BP; 34 A; 11 C; 24 G; 23 T;

Query Match 64.0%; Score 12.8; DB 1; Length 92;
 Best Local Similarity 87.5%; Pred. No. 5e+02;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 tttaaacatgcac 20
 |||||
 DB 14 TTTTAAAGCATCACAC 29

RESULT 11
 ID Q82182/c

```
ID 082182 standard; DNA; 22 BP.
AC 082182;
DE 04-SEP-1995 (first entry)
KW Chromosome 11 (locus D11S1075) STS primer CSR-349-tz.
KW sequence sampled mapping; genomic analysis; complex genome mapping;
KW cosmid library; chromosome 11; sequence tagged site; STS analysis; ss.
OS Synthetic.
PN W09429486-A.
PD 22-DEC-1994.
PE 15-JUN-1994; U06810.
PR 07-SEP-1993; US-078471.
PR 15-JUN-1993; US-117952.
PR (SALK ) SALK INST BIOLOGICAL STUDIES.
PI Evans GA, Smith MW;
DR WPI; 95-036508/05.
PT Sequencing complex genomes, present as fragments in a cosmid
PT library - by sequencing end-specific nucleotides of each clone
PT then correlating with spatial relationship of cosmid, esp. for
PT mammalian chromosomes.
PS Example 4; Page 69; 128pp; English.
CC Sequences were determined from the ends of chromosome 11-specific
CC cosmid by automated sequencing without intermediate subcloning.
CC A sample of 371 DNA sequence fragments were determined and of
CC these, 277 were suitable for STS primer prediction by computer.
CC analysis (using the "primer" program available from E. Lander, MIT).
CC The STSs and cosmid were mapped by in situ hybridisation, somatic
CC cell hybrid analysis or both. Using this method, 370 STSs specific
CC for human chromosome 11 were generated and most of them were
CC regionally mapped. This procedure illustrates a novel method for
CC sequencing complex genomes, designated "sequence sampled mapping".
CC The sequence sampled mapping method is useful for the completion of
CC high density sequence-based maps, and ultimately for the complete
CC sequencing of genomic DNA directly from cosmid clones.
CC See Q82001-082706 for STS primers.
SQ Sequence 22 BP; 6 A; 2 C; 8 G; 6 T;

Query Match 63.0%; Score 12.6; DB 1; Length 22;
Best Local Similarity 78.9%; Pred. No. 5.3e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ccccttaaacatcgcca 19
DB 22 CCCCTTAACATGCTCA 4

RESULT 12
082346/c
ID 082346 standard; DNA; 22 BP.
AC 082346;
DE 08-SEP-1995 (first entry)
KW Chromosome 11 (locus D11S1159) STS primer CSR-544-tz.
KW sequence sampled mapping; genomic analysis; complex genome mapping;
KW cosmid library; chromosome 11; sequence tagged site; STS analysis; ss.
OS Synthetic.
PN W09429486-A.
PD 22-DEC-1994.
PE 15-JUN-1994; U06810.
PR 15-JUN-1993; US-078471.
PR 07-SEP-1993; US-117952.
PR (SALK ) SALK INST BIOLOGICAL STUDIES.
PI Evans GA, Smith MW;
DR WPI; 95-036508/05.
PT Sequencing complex genomes, present as fragments in a cosmid
PT library - by sequencing end-specific nucleotides of each clone
PT then correlating with spatial relationship of cosmid, esp. for
PT mammalian chromosomes.
PS Example 4; Page 77; 128pp; English.
CC Sequences were determined from the ends of chromosome 11-specific
CC cosmid by automated sequencing without intermediate subcloning.
CC A sample of 371 DNA sequence fragments were determined and of
CC these, 277 were suitable for STS primer prediction by computer.
CC analysis (using the "primer" program available from E. Lander, MIT).
```

```
CC The STSs and cosmid were mapped by in situ hybridisation, somatic
CC cell hybrid analysis or both. Using this method, 370 STSs specific
CC for human chromosome 11 were generated and most of them were
CC regionally mapped. This procedure illustrates a novel method for
CC sequencing complex genomes, designated "sequence sampled mapping".
CC The sequence sampled mapping method is useful for the completion of
CC high density sequence-based maps, and ultimately for the complete
CC sequencing of genomic DNA directly from cosmid clones.
CC See Q82001-082706 for STS primers.
SQ Sequence 22 BP; 4 A; 4 C; 7 G; 7 T;

Query Match 63.0%; Score 12.6; DB 1; Length 22;
Best Local Similarity 78.9%; Pred. No. 5.3e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 ccccttaaacatcgccac 20
DB 21 CTCCTTACATAGCCAC 3

RESULT 13
T33684
ID T33684 standard; DNA; 33 BP.
AC T33684;
DE 14-JAN-1997 (first entry)
DE Primer used to amplify budR gene and budA promoter.
KW Promoter; regulatory protein; regulation; gene expression;
KW induction; inclusion bodies; ss.
OS Synthetic.
PN EP-733704-A2.
PD 25-SEP-1996.
PE 21-MAR-1996; 104475.
PR 24-MAR-1995; DE-010930.
PR 13-APR-1995; DE-014056.
PR (CONE ) CONSORTIUM ELEKTROCHEM IND GMBH.
PI Boeck A, Candussio A, Mayer D, Schlensof V;
DR WPI; 96-427054/43.
PT Protein expression system regulatable by acetate, pH or oxygen
PT includes regulatory protein and promoter from Klebsiella terrigena,
PT to provide simple method for controlling expression under all growth
PT conditions.
PS Example 10; Page 7; 41pp; German.
CC An expression system which is regulatable by acetate, pH and O2
CC comprises a regulatory protein (W02378) acting in trans and this
CC promoter. The expression system provides a simple and inexpensive
CC method for controlling gene expression under all growth conditions.
CC independently of the growth phase or the O2 content of the medium.
CC The use of three control variables allows a precisely staged
CC induction to be achieved and protein can be produced at a level that
CC does not result in the formation of inclusion bodies. Two primers
CC (T33682, T33684) were used to amplify a 1023 base pair fragment from
CC the plasmid pBAK16 contained the budR gene and the budA promoter.
CC The sequence was subcloned into the vector pBR32 and used to
CC generate a bud promoter/lacZ fusion to produce a reporter
CC construct.
SQ Sequence 33 BP; 11 A; 11 C; 7 G; 4 T;

Query Match 63.0%; Score 12.6; DB 1; Length 33;
Best Local Similarity 78.9%; Pred. No. 5.6e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ccccttaaacatcgcca 19
DB 7 CCCGCTTAACATCTCCA 25

RESULT 14
V09404/c
ID V09404 standard; DNA; 60 BP.
AC V09404;
DE 18-MAY-1998 (first entry)
```

DE Human ICE LAP-6 cDNA amplifying 3' primer 1.
KM ICE LAP-6; Interleukin-1 beta converting enzyme apoptosis protease-6;
KW treatment; viral infection; tumour; inflammation; osteoporosis;
KM AIDS; Alzheimer's disease; human; PCR primer; ss.
OS Synthetic.
OS Homo sapiens.
PN EP-808904-A2.
PD 26-NOV-1997.
PF 19-MAY-1997; 303397.
PR 05-JUN-1996; US-018961.
PR 20-MAY-1996; US-017949.
PR 23-MAY-1996; US-020344.
PA (HUMA-) HUMAN GENOME SCI INC.
PA (SMTK) SMITHKLINE BEECHAM CORP.
PA (UNMI) UNIV MICHIGAN.
PI Dixit VM, He W, Kikly KK, Ruben SM;
DR WPI; 98-001790/01.
PT DNA encoding interleukin-1 beta converting enzyme apoptosis
PT protease-6 - useful to develop products to treat, e.g. viral
PT infection, tumour, Alzheimer's disease, inflammation, osteoporosis
PT and AIDS
PT Example 1; Page 24; 44pp; English.
CC This primer is used for the PCR amplification of a human interleukin-1
CC beta converting enzyme apoptosis protease-6 (ICE LAP-6) cDNA. The ICE
CC LAP-6 polypeptide and agonists to the polypeptide can be used to induce
CC apoptosis, e.g. as an antiviral or antitumour agent, control embryonic
CC development and tissue homeostasis and the roles of such factors in
CC dysfunction and disease. Antagonists which inhibit the activity of the
CC ICE LAP-6 polypeptide can be used to treat, Alzheimer's or Parkinson's
CC disease, rheumatoid arthritis, septic shock, sepsis, stroke, chronic,
CC acute or central nervous system inflammation, osteoporosis, ischemia
CC reperfusion injury, cell death associated with cardiovascular disease,
CC polycystic kidney disease, degenerative liver disease, multiple sclerosis,
CC cardiovascular disease, degenerative liver disease, multiple sclerosis,
CC cerebellar degeneration, ischaemic injury, myocardial infarction,
CC acquired immunodeficiency syndrome (AIDS), myelodysplastic syndrome,
CC aplastic anaemia, male pattern baldness and head injury damage. They can
CC also be used for detection and diagnosis.
SQ Sequence 60 BP; 12 A; 4 C; 21 G; 23 T;

Query Match 62.0%; Score 12.4; DB 1; Length 60;
Best Local Similarity 92.9%; Pred. No. 7.6e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 ccttaaacatgc 17
|||||
DB 43 CTTTAAACATCAC 30

RESULT 15
X02116
ID X02116 standard; DNA: 28 BP.
AC X02116.
DT 23-APR-1999 (first entry)
DE Human FEN-1 PCR primer #2.
KW FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage; PCR primer; ss.
OS Synthetic.
OS Homo sapiens.
PN US3874283-A.
PD 23-FEB-1999.
PF 30-MAY-1995; 455968.
PR 30-MAY-1995; US-455968.
PA (HARR/) HARRINGTON J J.
PA (HSE/) HSEH C.
PA (LIEB/) LIEBER M R.
PI Harrington JJ, Hsieh C, Lieber MR;
DR WPI; 99-179985/15.
PT DNA encoding flap endonuclease polypeptides - useful for producing
PT e.g. recombinant polypeptides
PS Disclosure; Column 25; 58pp; English.
CC X02115-X02116 are PCR primers used in the isolation and amplification of

CC a human FEN-1 (flap endonuclease) protein. This protein can be used in
CC methods for detecting a pathological condition in a patient, for
CC diagnostic purposes, for screening for antineoplastic agents and
CC carcinogens, for diagnostic staging of neoplasia, for producing
CC recombinant flap endonuclease for use as research or diagnostic
CC reagents, for producing antibodies reactive with the novel polypeptides,
CC for producing transgenic nonhuman animals expressing the novel polypeptides
CC encoded by a transgene. The invention also provides novel molecular
CC cloning techniques and reagents involving cleavage of a flap or nick with
CC a flap endonuclease.
SQ Sequence 28 BP; 4 A; 9 C; 1 G; 14 T;

Query Match 61.0%; Score 12.2; DB 1; Length 28;
Best Local Similarity 82.4%; Pred. No. 8.8e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ccccttaaacatgc 17
|||||
DB 9 CCCCTTTTAAACTTCCC 25

Search completed: June 4, 2000, 16:24:14
Job time: 28910 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:19 ; Search time 4521.53 Seconds
(without alignments)
17.929 Million cell updates/sec

Title: US-09-164-714-14

Perfect score: 20

Sequence: 1 ccccttaaacatcgccac 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4857316 segs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *
44: gb_est25: *

45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est32: *
76: em_est33: *
77: em_est34: *
78: em_est35: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_est48: *
83: gb_est49: *
84: gb_est50: *
85: gb_est51: *
86: em_est36: *
87: em_est37: *
88: em_est38: *
89: em_est39: *
90: gb_est52: *
91: gb_est53: *
92: gb_est54: *
93: gb_est55: *
94: gb_est56: *
95: em_est40: *
96: em_est41: *
97: em_est42: *
98: em_est43: *
99: em_est44: *
100: em_est45: *
101: em_est46: *
102: gb_est57: *
103: gb_est58: *
104: em_est47: *
105: gb_est59: *
106: gb_est60: *
107: gb_est61: *
108: gb_est62: *
109: gb_est63: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

TITLE
JOURNAL
COMMENT

Mashu zebrafish EST Project 1998
Unpublished (1998)
On Jun 5, 1998 this sequence version replaced gI:3189655.
Other_ESTs: Fd03d07.xl
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: szbrafish@watson.wustl.edu
CDNA Library Preparation: Matthew Clark. CDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
ReSourceGen/ZentrumPrimatDatenbank, Berlin, Germany (web address:
www.fzpd.de)
Trace considered overall poor quality
Possible reversed clone; similarly on wrong strand
Seq primer: T3 ET from Amersham
High quality sequence stop: 1.

Location/Qualifiers

I . 70
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_lib="Zebrafish Mashu MPIMG EST"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield
stage embryos"
/lab_host="Xrl-blue MRF"
/note=Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
strand cDNA was primed with a Not I - oligo(dT)15 primer
[5'-pgcactgcttcagatccgcaggcgcccggccgttttttttttttttt-3'];
double-stranded cDNA was ligated to Sal I adaptors (BRL),
digested with Not I and cloned into the Not I and sal I
sites of the pSPORT vector (BRL). Library was constructed
by Matthew Clark (Lehrach lab, ICRF, London and Max Planck
Institut fuer Molekulare Genetik,Berlin). cDNAs for EST
analysis were selected following oligonucleotide
hybridization fingerprinting of arrayed clones from
zebrafish late somitogenesis (26 ss); adult liver or
embryonic shield stage (5.6 h) libraries. Fingerprint
data were used to computationally cluster cDNAs, and a
single cDNA from each cluster was chosen for sequencing.
In some cases multiple members of the same cluster were
sequenced to assess clustering parameters or single clones

BASE COUNT	19 a	11 c	20 g	20 t
ORIGIN				
Query Match		69.0%;	Score 13.8;	DB 63;
Best Local Similarity		88.2%;	Pred. No. 5.4e+03;	Length 70;
Matches 15;	Conservative	0;	Mismatches 2;	Indels 0;
Gaps 0;				
QY	4	ctttaaacatgcccac	20	
	111	11111111111111		
Db	64	CTTGAAACATCCCCAC	48	
RESULT 2				
AA985666/c				
LOCUS	AA985666	49 bp	mRNA	EST
DEFINITION	cr71c08.s1 NCLCGAP_Lu5 Homo sapiens cDNA clone IMAGE:1601295 3'			
	similar to SW:NDIM.HUMAN O00217 NADH-UBIQUINONE OXIDOREDUCTASE 23			
ACCESSION	AA985666			
VERSION	AA985666.1	GI:3167055		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 49)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL On Jan 14, 1998 this sequence version replaced gi:1798731.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bdrp/image/image.html

Trace considered overall poor quality
Insert Length: 517 Std Error: 0.00
Seq Primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES

source

1. 49
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1601295"
/clone.lib="NCI.CGAP.Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: Lung; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT

12 a 7 c 15 g 15 t

ORIGIN

Query Match 67.0%; Score 13.4; DB 40; Length 49;

Best Local Similarity 93.3%; Pred. No. 9.1e+03;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ccccttaaacatcg 16
Db 36 ccccttATACATCG 22

RESULT 3

A1423836/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 91)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)JOURNAL On Apr 21, 1998 this sequence version replaced gi:3071506.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLNL; contact the Image Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 68.

FEATURES

source

1. 91
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2110745"
/clone.lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT

28 a 15 c 18 g 30 t

ORIGIN

Query Match 67.0%; Score 13.4; DB 46; Length 91;

Best Local Similarity 93.3%; Pred. No. 8.4e+03;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 ttaaacaatcgccac 20
Db 58 TTTAAACATCGCCAC 44

RESULT 4

R03760

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Caenorhabditis briggsae

Eukaryota; Metazoa; Nematoda; Secernentea; Rhabdita; Rhabditidae; Rhabditina; Rhabditioidea; Rhabditidae; Pelodermidae; Caenorhabditis.

1 (bases 1 to 57)

Hiller, L., Chapell, B., Chisoe, S., Clark, N., Couch, J., Dubuque, T., Hawkins, M., Holman, M., Hultman, M., Kuwaba, T., Kuwabara, P., Le, M., Mardis, E., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevasakis, E., Waterston, R., Wohlmann, P. and Wilson, R.

Washington University Caenorhabditis briggsae EST project

Unpublished (1995)

Other ESTs: PK10H10.r1

Contact: Marra MA

Washington University Genome Sequencing Center

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1455

Fax: 314 286 1810

Email: mmarra@watson.wustl.edu

PCR_F: TGTAAACGACGCGCCAGTACAGTTACGCTCG

PCR_B: CAGAAACGCTATGACCTTATGAGATTTCTTCAGGTA

Source: Washington University Genome Sequencing Center

PCR amplified DNA is available from Washington University Genome Sequencing Center. Aliquots of the library may be requested from P. Kuwabara (pekemrc-lmb.cam.ac.uk).

Insert Length: 1138 Std Error: 0.00

Seq primer: Commercially available -21M13 dye primer

LOCUS AA730470 76 bp mRNA EST 06-JAN-1998
DEFINITION nm47612.s1 NCI-CGAP_Ew1 Homo sapiens CDNA clone IMAGE:1249774
similar to TR:Q16236 Q16236 NRP2. ; mRNA sequence.
ACCESSION AA730470
VERSION AA730470.1 GI:2751674
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 76)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1998 this sequence version replaced gl:1393484.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

FEATURES
source
1..76
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1249774"
/clone_1lb="NCI-CGAP_Ew1"
/tissue_type="Ewing's sarcoma"
/lab_host="DH10B"
/note="Vector: PAMP10; mRNA made from Ewing's sarcoma,
CDNA made by oligo-dT priming. Non-directionally cloned.
Size-selected on agarose gel, average insert size 600 bp.
Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."

BASE COUNT 27 a 16 c 15 g 18 t
ORIGIN

Query Match 65.0%; Score 13; DB 37; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 aaacatcgccac 20
|||||
Db 36 AAAACATCGCCAC 24

RESULT 8
LOCUS AI568312 70 bp mRNA EST 29-MAR-1999
DEFINITION tm68607.x1 NCI-CGAP_Lym12 Homo sapiens CDNA clone IMAGE:2174724 3'
repetitive element ; mRNA sequence.
ACCESSION AI568312
VERSION AI568312
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 70)
<http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
On Jun 22, 1998 this sequence version replaced gl:3247277.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Life Technologies catalog #: 11547-015
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

FEATURES
source
1..70
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2174724"
/clone_1lb="NCI-CGAP_Lym12"
/tissue_type="lymphoma, follicular mixed small and large
cell"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
SalI; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.25 kb. Life Technologies
catalog #: 11547-015"

BASE COUNT 19 a 13 c 32 g 6 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 48; Length 70;
Best Local Similarity 87.5%; Pred. No. 1.8e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ccccttaaacatcg 16
|||||
Db 41 CCCCTTAACAAACCG 26

RESULT 9
LOCUS AA844655/c 76 bp mRNA EST 31-DEC-1998
DEFINITION aj57d03.s1 Soares,Testis_NHT Homo sapiens CDNA clone IMAGE:1394405
3', mRNA sequence.
ACCESSION AA844655
VERSION AA844655
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 76)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 17, 1998 this sequence version replaced gl:2043793.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

Insert Length: 1467 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.
Location/Qualifiers

```
source
1. 76
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1394405"
/clone.lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5']
TGTTCACATCTGAGTGGAGCGCCGCCCAATTTTTTTTTT 3'}.
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to c0t5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
```

BASE COUNT 18 a 17 c 22 g 19 t

ORIGIN

Query Match 64.0%; Score 12.8; DB 39; Length 76;
Best Local Similarity 87.5%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ccccttaaacatcg 16
||||||| 11 |||||

Db 73 CCCCTTGAACCATCG 58

RESULT 10
AI347508 81 bp mRNA EST 08-APR-1999
LOCUS AI347508.1 NCI_CGAP_Col6 Homo sapiens cDNA clone IMAGE:2063100 3',
DEFINITION mRNA sequence.
ACCESSION AI347508.1 GI:4084714
VERSION AI347508.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 81)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

JOURNAL
COMMENT

FEATURES
source
Seq primer: -40bp from Gibco.
Location/Qualifiers
1. 81
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2063100"
/clone.lib="NCI_CGAP_Col6"
/tissue_type="Colon tumor, RER+"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI-CGAP_Col6 was
prepared, and ss circles were made in vitro. Following HAP

```
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1057416-1061255, and 114584-1145351).
Subtraction by Bento Soares and M. Fatima Bonaldo. "
```

BASE COUNT 18 a 21 c 22 g 20 t

ORIGIN

Query Match 64.0%; Score 12.8; DB 45; Length 81;
Best Local Similarity 87.5%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ccccttaaacatcg 16
||||||| 11 |||||

Db 76 CCCCTTGAACCATCG 61

RESULT 11
AA649968 90 bp mRNA EST 13-NOV-1997
LOCUS AA649968.1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:1187556 3',
DEFINITION mRNA sequence.
ACCESSION AA649968
VERSION AA649968.1 GI:2577296
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 90)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Nov 29, 1993 this sequence version replaced gi:430517.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

JOURNAL
COMMENT

FEATURES
source
Insert Length: 191 Std Error: 0.00
Seq primer: -40m13 fwd. EV from Amersham.
Location/Qualifiers
1. 90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1187556"
/clone.lib="NCI_CGAP_Pr22"
/sex="male"
/tissue_type="normal prostate"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; 1st strand cDNA was prepared
from normal prostate bulk tissue, and was then primed with
a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 23 a 22 c 24 g 21 t

ORIGIN

Query Match 64.0%; Score 12.8; DB 36; Length 90;
Best Local Similarity 87.5%; Pred. No. 1.7e+04;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ccccttaaacatcg 16
||||| ||| ||| |||
Db 75 CCCCTTGAACATCG 60

RESULT 12
H55177/c 92 bp DNA EST 07-DEC-1995
LOCUS CHR220116 Chromosome 22 exon Homo sapiens genomic clone C22_151 5',
DEFINITION mRNA sequence.
ACCESSION H55177
VERSION H55177.1 GI:1108043
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 92)
Trottauer,J.A., Long,K.R., Murrell,J.R., Stottler,C.J., Gusella,J.F.
and Buckler,A.J.
An expression-independent catalog of genes from human chromosome 22
Genome Res. 5 (3), 214-224 (1995)
96159527
On Apr 14, 1993 this sequence version replaced gi:639258.
CONTACT: Buckler AJ
Molecular Neurogenetics Unit
Massachusetts General Hospital
Building 149, 13th St., Charlestown MA 02129
Tel: 6177249616
Fax: 6177265736
Email: buckler@helix.mgh.harvard.edu
Insert Length: 2240 Std Error: 0.00
Seq primer: T3
High quality sequence stop: 217.
Location/Qualifiers
1..92
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="C22_151"
/clone_1lb="Chromosome 22 exon"
/lab_host="E. coli DH5a"
/note="Vector: pBluescriptIIKS+. Site.1: Sal I; Site.2:
Bam HI (destroyed); Exons were isolated from human
chromosome 22 specific cosmids using a modification of
the method of exon amplification (Proc. Natl. Acad. Sci.
USA 88:4005-4009, 1991). Amplified exons were digested
with Sal I and Bgl II and subsequently cloned into
pBluescriptIIKS+ at the Sal I and Bam HI sites."

BASE COUNT 22 a 17 c 24 g 29 t

ORIGIN

Query Match 64.0%; Score 12.8; DB 23; Length 92;
Best Local Similarity 87.5%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 cttaaacatcgcca 19
||||| ||| ||| |||
Db 92 CTTCAAAACATCTCCA 77

RESULT 13
A1480496 55 bp mRNA EST 09-MAR-1999
LOCUS vg34a03.x1 Soares_mammary_gland_NBMG Mus musculus cDNA clone
DEFINITION IMAGE:863212 3' similar to TR:064306 Q64306 PROLINE-RICH PROTEIN.
; mRNA sequence.
ACCESSION A1480496
VERSION A1480496.1 GI:4373722
KEYWORDS EST.
SOURCE house mouse.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 55)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/cgi/cgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jun 22, 1998 this sequence version replaced gi:3246745.
CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:507300
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
Trace considered overall poor quality
Possible reversed clone; similarity on wrong strand
High quality sequence stop: 1.
Location/Qualifiers
1..55
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:863212"
/clone_1lb="Soares_mammary_gland_NBMG"
/sex="male"
/tissue.type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT7T3P-Pac
(Pharmacia) with a modified polylinker; Site.1: Not I;
Site.2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTACCAATCGAAGTGGAGCGCGCGAATGCTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library.
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

BASE COUNT 17 a 22 c 12 g 3 t 1 others

ORIGIN

Query Match 63.0%; Score 12.6; DB 47; Length 55;
Best Local Similarity 75.0%; Pred. No. 2.3e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgccac 20
||||| ||| ||| ||| |||
Db 7 CCCCTTAATAAANCCCGC 26

RESULT 14
A0014474 84 bp mRNA EST 03-AUG-1998
LOCUS A0014474 Schizosaccharomyces pombe late log phase cDNA
DEFINITION Schizosaccharomyces pombe cDNA clone spc09915, mRNA sequence.
ACCESSION A0014474
VERSION A0014474.1 GI:3369265
KEYWORDS EST.
SOURCE fission yeast.
ORGANISM Schizosaccharomyces pombe
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
Schizosaccharomycetaceae; Schizosaccharomyces.
1 (bases 1 to 84)
Moriyomo,M. and Mita,K.
Identification of expressed sequence tags of Schizosaccharomyces
pombe
Unpublished (1998)
On Jan 9, 1998 this sequence version replaced gi:837494.

Contact: Mitsuoki Morimyo
Genome Research Group
National Institute of Radiological Sciences
9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
Email: morimyo@nirs.go.jp

FEATURES
source
location/Qualifiers

1. 84
/organism="Schizosaccharomyces pombe"
/strain="972"
/db_xref="taxon:4896"
/clone="spc09915"
/clone_lib="Schizosaccharomyces pombe late log phase cDNA"
/sex="h minus"
/note="Vector: M13mp19; The cDNA library of
Schizosaccharomyces pombe was prepared by cloning cDNA
into the SmaI site of M13mp19 DNA and the direction of DNA
sequences was not always from 5' to 3'. The cDNA data of
Schizosaccharomyces pombe are available for searching on
the World Wide Web. (URL, <http://www.nirs.go.jp>)"

BASE COUNT 29 a 16 c 9 g 29 t 1 others
ORIGIN

Query Match 63.0%; Score 12.6; DB 41; Length 84;
Best Local Similarity 75.0%; Pred. No. 2.2e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 ccccttaaacatgcgcac 20
| | | | | | | | | | | | | | | | | |
Db 10 CCCTTTNATACATCGCTAC 29

RESULT 15

AU014499 84 bp mRNA EST 03-AUG-1998
LOCUS
DEFINITION AU014499 Schizosaccharomyces pombe late log phase cDNA
ACCESSION AU014499
VERSION AU014499
KEYWORDS AU014499.1 GI:3369290
SOURCE EST.
ORGANISM fission yeast.
Schizosaccharomyces pombe
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
Schizosaccharomycetaceae; Schizosaccharomyces.

REFERENCE 1 (bases 1 to 84)
AUTHORS Morimyo,M. and Mita,K.
TITLE Identification of expressed sequence tags of Schizosaccharomyces
pombe

JOURNAL Unpublished (1998)
COMMENT On Jan 9, 1998 this sequence version replaced gi:937426.

Contact: Mitsuoki Morimyo
Genome Research Group
National Institute of Radiological Sciences
9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
Email: morimyo@nirs.go.jp.
Location/Qualifiers

1. 84
/organism="Schizosaccharomyces pombe"
/strain="972"
/db_xref="taxon:4896"
/clone="spc09947"
/clone_lib="Schizosaccharomyces pombe late log phase cDNA"
/sex="h minus"
/note="Vector: M13mp19; The cDNA library of
Schizosaccharomyces pombe was prepared by cloning cDNA
into the SmaI site of M13mp19 DNA and the direction of DNA
sequences was not always from 5' to 3'. The cDNA data of
Schizosaccharomyces pombe are available for searching on
the World Wide Web. (URL, <http://www.nirs.go.jp>)"

BASE COUNT 29 a 16 c 9 g 29 t 1 others
ORIGIN

Query Match 63.0%; Score 12.6; DB 41; Length 84;
Best Local Similarity 75.0%; Pred. No. 2.2e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy 1 ccccttaaacatgcgcac 20
| | | | | | | | | | | | | | | | | |
Db 10 CCCTTTNATACATCGCTAC 29

Search completed: June 4, 2000, 13:53:22
JOD time: 20946 sec

100,000

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:44 ; Search time 244.64 Seconds
(without alignments)
10.627 Million cell updates/sec

Title: US-09-164-714-14
20

Sequence: 1 ccccttaaacatcgccac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 08
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/6.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/PCUS.COMB.seq:*
7: /cgn2_6/ptodata/2/1na/Backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.2	71.0	50	US-08-171-389-545	Sequence 545, App
2	14.2	71.0	50	US-08-123-936-545	Sequence 545, App
3	14.2	71.0	50	US-08-475-228A-545	Sequence 545, App
4	14.2	71.0	50	US-08-482-080A-545	Sequence 545, App
5	14.2	71.0	50	PCF-US93-12388-545	Sequence 545, App
6	12.6	63.0	22	US-08-117-952-182	Sequence 182, App
7	12.6	63.0	22	US-08-117-952-346	Sequence 346, App
8	12.6	63.0	22	US-08-614-686A-9	Sequence 9, App1
9	12.4	62.0	60	US-08-852-936C-6	Sequence 6, App1
10	12.2	61.0	28	US-08-455-968E-14	Sequence 14, App1
11	12.2	61.0	46	US-08-171-389-308	Sequence 308, App
12	12.2	61.0	46	US-08-123-936-308	Sequence 308, App
13	12.2	61.0	46	US-08-475-228A-308	Sequence 308, App
14	12.2	61.0	46	US-08-482-080A-308	Sequence 308, App
15	12.2	61.0	46	PCF-US93-12388-308	Sequence 308, App
16	12.2	60.0	49	US-08-045-806-22	Sequence 22, App1
17	12.2	60.0	49	US-08-366-051B-22	Sequence 22, App1
18	12.2	60.0	50	US-08-171-389-538	Sequence 538, App
19	12.2	60.0	50	US-08-123-936-538	Sequence 538, App
20	12.2	60.0	50	US-08-475-228A-538	Sequence 538, App
21	12.2	60.0	50	US-08-482-080A-538	Sequence 538, App
22	12.2	60.0	50	PCF-US93-12388-538	Sequence 538, App
23	12.2	60.0	66	US-08-852-936C-9	Sequence 9, App1
24	12.2	60.0	80	US-08-276-452A-64	Sequence 64, App1
25	12.2	60.0	80	US-08-798-744-64	Sequence 64, App1
26	11.8	59.0	20	US-08-609-572-6	Sequence 6, App1
27	11.8	59.0	23	US-08-179-557-12	Sequence 12, App1

28	11.8	59.0	89	2	US-08-162-102C-48	Sequence 48, App1
29	11.6	58.0	20	3	US-08-117-952-70	Sequence 70, App1
30	11.6	58.0	21	3	US-07-952-277A-82	Sequence 82, App1
31	11.6	58.0	48	3	US-08-642-045B-21	Sequence 21, App1
32	11.6	58.0	60	1	US-08-478-039-6	Sequence 6, App1
33	11.6	58.0	60	2	US-08-476-349A-6	Sequence 6, App1
34	11.6	58.0	74	4	US-08-708-856A-11	Sequence 11, App1
35	11.6	58.0	74	5	US-09-287-375-11	Sequence 11, App1
36	11.6	58.0	75	5	US-08-828-712-3	Sequence 3, App1
37	11.6	58.0	78	7	5171680-7	Patent No. 5171680
38	11.4	57.0	15	2	US-08-334-847-560	Sequence 560, App
39	11.4	57.0	46	2	US-08-462-388-8	Sequence 8, App1
40	11.4	57.0	48	3	US-08-811-492-124	Sequence 124, App
41	11.2	56.0	34	1	US-08-151-574-37	Sequence 37, App1
42	11.2	56.0	34	1	US-08-146-422-4	Sequence 4, App1
43	11.2	56.0	34	1	US-08-146-424-4	Sequence 4, App1
44	11.2	56.0	34	2	US-08-626-554-20	Sequence 20, App1
45	11.2	56.0	34	3	US-08-419-448-37	Sequence 37, App1

ALIGNMENTS

RESULT 1
US-08-171-389-545/C
Sequence 545, Application US/08171389
Patent No. 5578444
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 641
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/171,389
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993.
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175/G19P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 545:
SEQUENCE CHARACTERISTICS:

LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Epstein Barr virus R2 (start site)
US-08-171-389-545

Query Match 71.0%; Score 14.2; DB 1; Length 50;
Best Local Similarity 84.2%; Pred. No. 49;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccccttaaacatcgcca 19
||||| ||| ||| |||
Db 28 CCCCTTTAAATATCCCA 10

RESULT 2

US-08-123-936-545/c
; Sequence 545 Application US/08123936
; Patent No. 5726014
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; TITLE OF INVENTION: Screening Assay for the Detection of
; TITLE OF INVENTION: DNA-Binding Molecules
; NUMBER OF SEQUENCES: 640
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,936
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 545:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Epstein Barr virus R2 (start site)
; INDIVIDUAL ISOLATE: 88897)

US-08-123-936-545

Query Match 71.0%; Score 14.2; DB 2; Length 50;
Best Local Similarity 84.2%; Pred. No. 49;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccccttaaacatcgcca 19
||||| ||| ||| |||
Db 28 CCCCTTTAAATATCCCA 10

RESULT 3

US-08-475-228A-545/c
; Sequence 545 Application US/08475228A
; Patent No. 5869241
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk E.
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 664
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,228A
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081,070
; FILING DATE: 22-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stratford, Carol A.
; REGISTRATION NUMBER: 34,444
; REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 545:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Epstein Barr virus R2 (start site)
; INDIVIDUAL ISOLATE: 88897)
US-08-475-228A-545

Query Match 71.0%; Score 14.2; DB 3; Length 50;
Best Local Similarity 84.2%; Pred. No. 49;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgcca 19
||||| ||| ||| |||
DB 28 CCCCTTTAAATATCCCA 10

RESULT 4
US-08-482-080A-545/C
; Sequence 545, Application US/08482080A
; Patent No. 6010849
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk E.
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 664
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,080A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/171,389
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081,070
; FILING DATE: 22-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Brady, John F.
; REGISTRATION NUMBER: 39,118
; REFERENCE/DOCKET NUMBER: 4600-0175.20/619P3D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ. ID NO: 545:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Epstein Barr virus R2 (start site
; INDIVIDUAL ISOLATE: 88897)
US-08-482-080A-545

Query Match 71.0%; Score 14.2; DB 5; Length 50;
Best Local Similarity 84.2%; Pred. No. 49;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgcca 19
||||| ||| ||| |||
DB 28 CCCCTTTAAATATCCCA 10

RESULT 5
PCT-US93-12388-545/C
; Sequence 545, Application PC/TUS9312388
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 641
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12388
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0175.41/619PCT2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ. ID NO: 545:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Epstein Barr virus R2 (start site
; INDIVIDUAL ISOLATE: 88897)
PCT-US93-12388-545

Query Match 71.0%; Score 14.2; DB 6; Length 50;
Best Local Similarity 84.2%; Pred. No. 49;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgcca 19
||||| ||| ||| |||
DB 28 CCCCTTTAAATATCCCA 10

RESULT 6
US-08-117-952-182/C
; Sequence 182, Application US/08117952
; Patent No. 5851760

GENERAL INFORMATION:
APPLICANT: Evans, Glen A.
APPLICANT: Smith, Michael W.
TITLE OF INVENTION: METHOD FOR GENERATION OF SEQUENCE
TITLE OF INVENTION: SAMPLED MAPS OF COMPLEX GENOMES
NUMBER OF SEQUENCES: 797
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/117,952
FILING DATE: 07-SEP-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/078,471
FILING DATE: 15-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9423
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 182:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Oligonucleotide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-117-952-182

Query Match 63.0%; Score 12.6; DB 3; Length 22;
Best Local Similarity 78.9%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgcca 19
||||| ||||| |||
Db 22 CCCCTTCAAAACAGTGCCTA 4

RESULT 7
US-08-117-952-346/c
Sequence 346, Application US/08117952
Patent No. 5851760
GENERAL INFORMATION:
APPLICANT: Evans, Glen A.
APPLICANT: Smith, Michael W.
TITLE OF INVENTION: METHOD FOR GENERATION OF SEQUENCE
TITLE OF INVENTION: SAMPLED MAPS OF COMPLEX GENOMES
NUMBER OF SEQUENCES: 797
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/117,952
FILING DATE: 07-SEP-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/078,471
FILING DATE: 15-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9423
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 346:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Oligonucleotide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-117-952-346

Query Match 63.0%; Score 12.6; DB 3; Length 22;
Best Local Similarity 78.9%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ccccttaaacatcgccac 20
||| ||| ||||| |||||
Db 21 CTCCTTAGAACATAGCCAC 3

RESULT 8
US-08-614-686A-9
Sequence 9, Application US/08614686A
Patent No. 5830692
GENERAL INFORMATION:
APPLICANT: B CK, August; MAYER, Dagmar; SCHLENSOG,
TITLE OF INVENTION: Express System Which Can Be
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Collard & Roe, P.C.
STREET: 1077 No. 5830692thern Boulevard
CITY: Roslyn
STATE: New York
COUNTRY: U.S.A.
ZIP: 11576
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect Version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/614,686A
FILING DATE: MARCH 12, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GRI9510930
FILING DATE: 24 MARCH 1995
APPLICATION NUMBER: GRI9514056.7
FILING DATE: 13 APRIL 1995
ATTORNEY/AGENT INFORMATION:
NAME: Collard, Allison C.
REGISTRATION NUMBER: 22,532
REFERENCE/DOCKET NUMBER: BOCK ET AL.-4
ATTORNEY/AGENT INFORMATION:
NAME: Freedman, Edward R.
REGISTRATION NUMBER: 26,048

REFERENCE/DOCKET NUMBER: BOCK ET AL.-4
ATTORNEY/AGENT INFORMATION:
NAME: Keusey, Edwin H.
REGISTRATION NUMBER: 34,361
REFERENCE/DOCKET NUMBER: BOCK ET AL.-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 365-9802
TELEFAX: (516) 365-9805
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleotide
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: additional nucleic acid
DESCRIPTION: synthetic DNA
IMMEDIATE SOURCE:
CLONE: Oligo3
POSITION IN GENOME:
UNITS: bp
US-08-614-686A-9

Query Match 63.0%; Score 12.6; DB 3; Length 33;
Best Local Similarity 78.9%; Pred. No. 3.1e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ccccttaaacatcgcca 19
||| ||||| |||
Db 7 CCCGCTAAACATCTCCA 25

RESULT 9
US-08-852-936C-6/C
Sequence 6, Application US/08852936C
Patent No. 6010878
GENERAL INFORMATION:
APPLICANT: DIXIT, VISHVA M.
APPLICANT: HE, WEI-WU
APPLICANT: KIKLY, KRISTINE K.
APPLICANT: RUBEN, STEVEN M.
TITLE OF INVENTION: INTERLEUKIN-1 BETA CONVERTING
TITLE OF INVENTION: ENZYME LIKE APOPTOTIC PROTEASE-6
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & Prestia
STREET: P.O. Box 980
CITY: Valley Forge
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,936C
FILING DATE: 08-MAY-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,961
FILING DATE: 05-JUN-1996
APPLICATION NUMBER: 60/020,344
FILING DATE: 23-MAY-1996
APPLICATION NUMBER: 60/017,949
FILING DATE: 20-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: P50483-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700

TELEFAX: 610-407-0700
TELEX: 846169
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-852-936C-6

Query Match 62.0%; Score 12.4; DB 5; Length 60;
Best Local Similarity 92.9%; Pred. No. 4.2e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ctttaaacatcg 17
||||||| |
Db 43 CTTTAAACATCACC 30

RESULT 10
US-08-455-968E-14
Sequence 14, Application US/08455968E
Patent No. 5874283
GENERAL INFORMATION:
APPLICANT: Harrington, John L.
APPLICANT: Hsieh, Chih-Lin
APPLICANT: Lieber, Michael
TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 18985-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (oligonucleotide)
US-08-455-968E-14

Query Match 61.0%; Score 12.2; DB 3; Length 28;
Best Local Similarity 82.4%; Pred. No. 4.9e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccccttaaacatcg 17
||||| ||| |
Db 9 CCCCTTTAAACTCC 25

```
RESULT 11
US-08-171-389-308
; Sequence 308, Application US/08171389
; Patent No. 5578444
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk E.
; TITLE OF INVENTION: Sequence-directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 641
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/171.389
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123.936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081,070
; FILING DATE: 22-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0175/G19P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 308:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human thrombospondin gene
US-08-171-389-308

Query Match 61.0%; Score 12.2; DB 1; Length 46;
Best Local Similarity 82.4%; Pred. No. 5.2e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
; Patent No. 5726014
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk E.
; TITLE OF INVENTION: Screening Assay for the Detection of
; TITLE OF INVENTION: DNA-Binding Molecules
; NUMBER OF SEQUENCES: 640
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,936
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 308:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human thrombospondin gene
US-08-123-936-308

Query Match 61.0%; Score 12.2; DB 2; Length 46;
Best Local Similarity 82.4%; Pred. No. 5.2e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 ccccttaaacatgc 17
|||||
Db 17 CCCCTTAAAGCGCGC 33

RESULT 12
US-08-123-936-308
; Sequence 308, Application US/08123936
```

```
RESULT 13
US-08-475-228A-308
; Sequence 308, Application US/08475228A
; Patent No. 5869241
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk E.
; TITLE OF INVENTION: Sequence-directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 664
```



```

CORRESPONDENCE ADDRESS:
ADDRESS: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,228A
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stratford, Carol A.
REGISTRATION NUMBER: 34,444
REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 308:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human thrombospondin gene
US-08-475-228A-308

Query Match          61.0%; Score 12.2; DB 3; Length 46;
Best Local Similarity 82.4%; Pred. No. 5.2e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgc 17
Db 17 CCCCTTTAAAGCGCGC 33

RESULT 14
US-08-482-080A-308
Sequence 308, Application US/08482080A
Patent No. 6010849
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:
ADDRESS: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
```

```

STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,080A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/171,389
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Brady, John F.
REGISTRATION NUMBER: 39,118
REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 308:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human thrombospondin gene
US-08-482-080A-308

Query Match          61.0%; Score 12.2; DB 5; Length 46;
Best Local Similarity 82.4%; Pred. No. 5.2e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccccttaaacatcgc 17
Db 17 CCCCTTTAAAGCGCGC 33

RESULT 15
PCT-US93-12388-308
Sequence 308, Application PC/TUS9312388
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Sequence-directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 641
CORRESPONDENCE ADDRESS:
ADDRESS: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12388
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Radian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175.41/G19PCT2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 308:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human thrombospondin gene
PCT-US93-12388-308

Query Match 61.0%; Score 12.2; DB 6; Length 46;
Best Local Similarity 82.4%; Pred. NO. 5.2e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0;

OY 1 ccccttaaacatgc 17
|||||
Db 17 CCCCTTTAAAGCGCGC 33

Search completed: June 4, 2000, 16:09:45
Job time: 28064 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:14 ; Search time 322.35 Seconds
(without alignments)
34,927 Million cell updates/sec

Title: US-09-164-714-15

Perfect score: 45

Sequence: 1 ggcagcagatgcagcttaa.....ctatggcagatcagcttc 45

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 08
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.4	40.9	51	1	TSSR-1 mutagenic p
2	18	40.0	66	1	Sequence of C. tra
3	18	40.0	90	1	Sequence of Chalm
4	17.8	39.6	35	1	PCR primer set MPQ
5	17.8	39.6	35	1	Trichosanthin prim
6	17.8	39.6	47	1	PCR primer FL for
7	17.8	39.6	47	1	light chain primer
8	17.8	39.6	54	1	Sequence encoding
9	17.6	39.1	100	1	Human biallelic po
10	17.4	38.7	83	1	Orphan nuclear hor
11	17.4	38.7	83	1	L57 consensus nucl
12	16.8	37.3	86	1	Oligonucleotide V6
13	16.6	36.9	28	1	Primer used in iso
14	16.6	36.9	47	1	Forward primer for
15	16.6	36.9	47	1	V33995
16	16.6	36.9	80	1	Human gene signatu
17	16.4	36.4	50	1	Human 70 kda heat
18	16.4	36.4	50	1	Human 70 kda heat
19	16.4	36.4	50	1	Test sequence from
20	16.4	36.4	78	1	Canv355 promoter B
21	16.4	36.4	100	1	Candida albicans-s
22	16.2	36.0	69	1	Staphylococcus aur
23	16.2	36.0	81	1	Staphylococcus aur
24	16.2	36.0	81	1	Mouse orpnan hormo
25	15.8	35.1	29	1	Mouse VL30 retrotr
26	15.8	35.1	37	1	Amplification Prim
27	15.8	35.1	38	1	primer ABI40 for I
28	15.8	35.1	66	1	Sequence of C. tra
29	15.8	35.1	66	1	C. trachomatis ser
30	15.6	34.7	38	1	Probe B(1) for bov
31	15.6	34.7	38	1	Probe B(II) for bo
32	15.6	34.7	51	1	5mer synthesised
33	15.6	34.7	88	1	5'-3' oligo 3 for
34	15.6	34.7	90	1	5' leader sequence

c	35	15.4	34.2	41	1	034160	Sequence upstream
c	36	15.4	34.2	41	1	V81550	Oligonucleotide us
c	37	15.4	34.2	44	1	V81549	Oligonucleotide us
c	38	15.4	34.2	50	1	069753	Human cytomegalovir
c	39	15.4	34.2	50	1	T64215	HCWV 2.7 kb transc
c	40	15.4	34.2	50	1	X17503	Test sequence from
c	41	15.4	34.2	51	1	072798	TSSR-1 mutagenic p
c	42	15.4	34.2	87	1	035813	Annealing oligonuc
c	43	15.4	34.2	87	1	035812	Annealing oligonuc
c	44	15.4	34.2	87	1	T04761	PCR primer, HIVTM1
c	45	15.4	34.2	87	1	T04762	PCR primer, HIVTM2

ALIGNMENTS

RESULT	1	
ID	072792	standard; DNA; 51 BP.
AC	072792;	
DT	08-MAY-1995	(first entry)
DE	TSSR-1 mutagenic primer 3top.	
KW	Kawasaki syndrome; mucocutaneous lymph node syndrome;	
KW	toxic shock syndrome toxin-1; TSSR-1; superantigen; antigen;	
KW	Staphylococcus aureus; T-lymphocyte; immune response; vaccine;	
KW	diagnosis; primer; PCR; polymerase chain reaction; mutagenesis; ss.	
OS	Synthetic.	
PN	W09422474-A.	
PD	13-OCT-1994.	
PF	05-APR-1994; U03719.	
PR	05-APR-1993; US-042876.	
PA	(NADE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.	
PI	Kappler J, Kotzin B, Leung D, Marrack P;	
DR	WPI; 94-332822/41.	
PT	Treating or preventing Kawasaki Syndrome - by administering a	
PT	superantigen or deriv. to modify a T cell response or provoke an	
PT	immune response	
PS	Disclosure; Fig. 5; 50pp; English.	
CC	The gene for S. aureus TSSR-1 superantigen was isolated by PCR	
CC	amplification of genomic DNA using primers Atop (072788) and Btop	
CC	(072804) and Bbot (072805). Mutant TSSR-1 molecules were	
CC	synthesized using the mutagenic PCR primers given in 072789 and	
CC	072791-803. The method introduced random mutations in approx. 17	
CC	residues in the mature TSSR-1 protein. Selected mutated	
CC	superantigens selectively stimulate only some of the T-cell	
CC	populations stimulated by the wild-type superantigen.	
SQ	Sequence 51 BP; 18 A; 9 C; 9 G; 15 T;	
Query Match	40.9%;	Score 18.4; DB 1; Length 51;
Best local Similarity	78.6%;	Pred. No. 1.5e+02;
Matches 22; Conservative	0;	Mismatches 6; Indels 0; Gaps 0;
QY	17 ttaaaacattatgcagatcagctt 44	
DB	14 TAAAAAACACAGATGCGATCAGCCT 41	
RESULT	2	
ID	N97082/C	
AC	N97082;	
DT	06-MAR-1992	(first entry)
DE	Sequence of C. trachomatis serovar A major outer membrane protein (MOMP)	
DE	variable domain (VD) gene A-VDII base pairs 487-552	
KW	Chlamydia trachomatis; antigen; monoclonal antibody; vaccine;	
KW	diagnosis; serotyping; non-immunologic assay; ss.	
OS	Chlamydia trachomatis.	
FT	key	
FT	1. 66	
FT	Location/Qualifiers	
FT	/*tag= a	
PD	US7324664-A.	
PD	29-AUG-1989.	


```

FT      /mod_base= inosine
FT      18
FT      /tag- c
FT      /mod_base= inosine
FT      30
FT      /tag- d
FT      /mod_base= inosine
PN      US516056-A.
PD      24-NOV-1992.
PE      04-APR-1989; 333184.
PR      04-APR-1989; US-333184.
PR      07-SEP-1989; US-404326.
PR      09-DEC-1991; US-804293.
PA      (GENE-) GENELABS INC.
PI      Chow TP, Piatlak M;
PI      WPI; 92-414954/50.
PT      Recombinant tricosanthin protein prodn. in E. coli - for use in
PT      the selective inhibition of viral expression in HIV infected
PT      cells
PS      disclosure: Page 30; 37pp; English.
CC      Primer set MPQ-1 was designed for binding to the anti-sense strand
CC      of the TCS coding region.
CC      The TCS coding sequence was amplified using the primers of Q31828-30.
CC      The amplified prod. has the sequence of Q31827, which was used as a
CC      probe. One isolate, PQ21D, comprises the sequence of Q31826.
CC      The recombinant TCS sequence may be used in the recombinant prodn.
CC      of TCS. TCS can be used for the selective inhibition of viral
CC      expression in HIV-infected human T-cells or macrophages.
SQ      Sequence 35 BP; 8 A; 3 C; 8 G; 5 T;

Query Match
Best Local Similarity 39.6%; Score 17.8; DB 1; Length 35;
Matches 19; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY      7 catggcaacttaaacactatggcagatc 39
DB      33 CATNGCCTCTTTRANACRTATTTGCGCYTC 1

RESULT 6
Q24660
ID      Q24660 standard; DNA; 47 BP.
AC      Q24660;
DT      10-NOV-1992 (first entry)
DE      PCR primer FL for CAMPATH-1H light chain.
KW      Polymerase chain reaction; humanised antibody; CAMPATH-1H;
KW      rat anti-human CD18 light chain; YFC51.1.1;
KW      human IgG1 heavy chain; PCR grafting; ss.
OS      Synthetic.
PN      WO9207075-A.
PD      30-APR-1992.
PR      08-OCT-1991; G01744.
PR      10-OCT-1990; GB-022011.
PA      (WEILL) WELLCOME FOUNDD LTD.
PI      Crowe JS, Lewis AP;
PI      WPI; 92-167155/20.
PT      Prep. of chimeric humanised antibodies - using a new polymerase
PT      chain reaction technique
PS      Example 2; Page 45; 67pp; English.
PS      The YFC51.1.1 rat anti-human -CD18 light chain was humanised as
CC      follows:
CC      Primer EL (Q24659) was used with Primer FL (Q24660) in a PCR
CC      reaction using as template CAMPATH-1H light chain (i.e. humanised
CC      CAMPATH-1 on REI framework; Biotechnology 9:64-68 (1991)) to
CC      produce fragment EFL. Three other PCR reactions were performed on
CC      the same template, generating fragments ABL, CDL and GHL. Fragments
CC      EFL and GHL were combined and used as the template for a PCR
CC      reaction with primers EL and HL (Q24662) to produce fragment EHL.
CC      Similarly, fragment ADL was produced from ABL and CDL using the
CC      primers AL and DL (Q24655 and Q24658, respectively). The products
CC      ADL and EHL were purified and combined in a recombinant PCR
CC      reaction using primers AL and HL. The final humanised light chain

```

```

CC      product, AHL, was cloned into the HindIII site of pUC18 (primers
CC      AL and HL both contain HindIII sites).
SQ      Sequence 47 BP; 9 A; 7 C; 16 G; 15 T;

Query Match
Best Local Similarity 39.6%; Score 17.8; DB 1; Length 47;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      15 cttaaaacactatggcagatcagct 43
DB      11 CTTTCATATACCTGTGCGAGTAGTAGT 39

RESULT 7
Q35185
ID      Q35185 standard; cDNA; 47 BP.
AC      Q35185;
DT      08-JUN-1993 (first entry)
DE      Light chain primer FL.
KW      Probe; myeloma; Y3-Ag 1.2.3; primer; rat; monoclonal; antibody; COS;
KW      YFC51.1.1; CD18; humanised; antigen; leukocyte; lung; sepsis; asthma;
KW      endotoxic shock; adult respiratory distress syndrome; inflammation;
KW      immunotoxin; transient expression; PCR; polymerase chain reaction; ss.
OS      Synthetic.
PN      WO9302191-A.
PD      04-FEB-1993.
PE      15-JUL-1992; G01289.
PR      16-JUL-1991; GB-015364.
PR      (WEILL) WELLCOME FOUNDD LTD.
PI      Crowe S, Sims M, Waldmann H;
PI      WPI; 93-058788/07.
PT      New humanised antibody specific for human CD-18 antigen -
PT      inhibits influx of leukocytes into the lungs, useful for treating
PT      endotoxic shock, adult respiratory distress syndrome, asthma, etc.
PS      disclosure: Page 48; 59pp; English.
CC      The sequences given in Q35180-87 are primers which were used to
CC      amplify and humanise the light chain isolated from the rat antibody
CC      YFC51.1.1. The light chain of YFC51.1.1 was isolated using a non-
CC      radioactively labelled clone of the light chain from rat myeloma
CC      Y3-Ag 1.2.3. The isolated sequences were amplified, humanised and
CC      constructed into the light chain genes using these primers. The gene
CC      construction, and a corresponding one for the heavy chain (see also
CC      Q35186-95) were transformed into COS cells which transiently
CC      expressed the humanised YFC51.1.1. YFC51.1.1 is a CD18 antibody
CC      which was used as a basis for the production of a humanised antibody
CC      with specificity for CD18 antigen. The antibody may be useful in
CC      treating leukocyte-mediated conditions, such as inhibiting influx of
CC      leukocytes into the lung and other organs during sepsis, endotoxic
CC      shock or adult respiratory distress syndrome. The antibodies may
CC      also be used to treat asthma and inflammation and may form part of
CC      an immunotoxin.
SQ      Sequence 47 BP; 9 A; 7 C; 16 G; 15 T;

Query Match
Best Local Similarity 39.6%; Score 17.8; DB 1; Length 47;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      15 cttaaaacactatggcagatcagct 43
DB      11 CTTTCATATACCTGTGCGAGTAGTAGT 39

RESULT 8
NB1348/C
ID      NB1348 standard; DNA; 54 BP.
AC      NB1348;
DT      15-NOV-1990 (first entry)
DE      Sequence encoding modified AA sequence of signal peptide for
DE      egg white lysozyme
KW      Modified egg white lysozyme signal sequence; enzyme; ds.
OS      Homo sapiens.

```

EH Key Location/Qualifiers
FT signal_peptide 1..54
/tag= a
PN J63233789-A.
PD 29-SEP-1988.
PF 23-MAR-1987; 069764.
PR 23-MAR-1987; JP-069764.
PA (AUIJN) Tampaku Kagaku Kenk.
DR WPI; 88-318076/45.
P-PSDB; P81021.
PT DNA sequence encoding signal peptide -
PT having good protein secretion activity
PS Claim 4; Page 523; 11pp; Japanese.
CC AAs of egg white lysozyme other than 1st, 2nd and 16th-18th residue
CC are all or mostly substituted with one kind of hydrophobic AA. Using the
CC modified signal peptide, foreign proteins, esp. human lysozyme, can be
CC secreted correctly and at a high level.
SQ Sequence 54 BP; 6 A; 5 C; 16 G; 27 T;

Query Match 39.6%; Score 17.8; DB 1; Length 54;
Best Local Similarity 62.2%; Pred. No. 2.5e+02;
Matches 28; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 ggaagcattgcaacttaaaacactattggcagatcagcttc 45
| | | | | | | | | | | | | | | | | | | | | |
DB 47 GCAGCGCAATGCAACATTAACAACACATATGCAACACATCTC 3

RESULT 9
X10693
ID X10693 standard; DNA; 100 BP.
AC X10693;
DE 30-MAR-1999 (first entry)
KW Human biallelic polymorphic DNA fragment WI-9558.
KW Polymorphism: biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PI (MHED) WHITEHEAD INST BIOMEDICAL RES.
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 72; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Leech-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberosus sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 33 A; 24 C; 15 G; 27 T;

Query Match 39.1%; Score 17.6; DB 1; Length 100;
Best Local Similarity 71.9%; Pred. No. 3.2e+02;
Matches 23; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 9 tggaactttaaaacactattggcagatc 40
| | | | | | | | | | | | | | | | | | | | | |
DB 24 TGCTATCTTTAGAACACTTTTCAGCAAGATCA 55

RESULT 10
T14004/c
ID T14004 standard; CDNA; 83 BP.
AC T14004;
DE 19-JUN-1996 (first entry)
KW Orphan nuclear hormone receptor clone L57 (nt233-315).
KW Obesity; body weight disorder; diagnosis; therapy; appetite; fat;
KW gene modulator; adipose; orphan nuclear hormone receptor; ss.
OS Mus musculus.
PN WO9605861-A1.
PD 29-FEB-1996.
PF 23-AUG-1995; U10918.
PR 23-AUG-1994; US-294522.
PR 06-JUN-1995; US-470868.
PA (MILL-) MILLENIUM PHARM INC.
PI Tartaglia LA;
DR WPI; 96-151150/15.
PT New gene which is differentially expressed in body weight disorders
PT - corresp. gene prod., antibodies and gene modulators; for
PT diagnosing and treating a weight disorder, such as obesity
PS Example 9; Fig 13; 162pp; English.
CC A set point paradigm was used to identify mouse gene sequences
CC which are differentially expressed and which may contribute to body
CC wt. disorders and/or may be involved in body wt. regulation or
CC appetite modulation. cDNA clone L57 was isolated. Nucleotides
CC 39-150 (T14002), 151-269 (T14003), 233-315 (T14004) and 4-41
CC (T14005) correspond to nucleotides 1321-1432, 1432-1550, 1513-1595
CC and 1285-1322, respectively, of the mouse orphan nuclear hormone
CC receptor. L57 gene expression is higher in the liver of
CC underweight mice than in that of overweight littermates. L57,
CC and other identified genes (see T13979-T14001 and T14006-T14008),
CC can be used to diagnose body wt. disorders.
SQ Sequence 83 BP; 29 A; 21 C; 8 G; 24 T;

Query Match 38.7%; Score 17.4; DB 1; Length 83;
Best Local Similarity 64.9%; Pred. No. 3.7e+02;
Matches 24; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 2 ggaagcattgcaacttaaaacactattggcagat 38
| | | | | | | | | | | | | | | | | | | | | |
DB 61 GAAGAYCTGCAACCTTACAGACTATTTGTAAATAAT 25

RESULT 11
X03753/c
ID X03753 standard; DNA; 83 BP.
AC X03753;
DE 31-MAR-1999 (first entry)
KW L57 consensus nucleotide sequence #3.
KW Body weight disorder; obesity; appetite regulation; thermoregulation;
KW anorexia; abnormal food intake; cachexia; thermogenesis; ss.
OS Synthetic.
PN US5853975-A.
PD 29-DEC-1998.
PF 26-FEB-1997; 807861.
PR 26-FEB-1997; US-807861.
PR 23-AUG-1994; US-294522.
PR 06-JUN-1995; US-470868.
PR 23-AUG-1995; US-518878.
PA (MILL-) MILLENIUM PHARM INC.
PI Tartaglia LA;
DR WPI; 99-094892/08.
PT Identifying modulators of c5 protein uncoupling activity - used as

potential regulators of thermogenesis, appetite and body weight,
 e.g. for treating obesity, anorexia and cachexia
 Example: Fig 13; 83bp; English.
 The present invention describes a method for identifying a compound (1)
 that modulates C5 protein uncoupling activity by: (a) treating a test
 compound with C5 or cells expressing it; (b) measuring the level of
 uncoupling activity; and (c) comparing this with level of activity in
 the absence of the test compound. Any difference in the levels indicates
 a modulator. C5 protein has uncoupling activity, i.e. it can transport
 protons across the mitochondrial inner membrane, reducing the proton
 motive force and allowing dissipation of caloric energy as heat. It is
 thus a regulator of thermogenesis and is involved in body weight
 regulation. (1) are potentially useful in treatment of body weight
 disorders, regulation of appetite and thermoregulation, e.g. in cases of
 obesity, anorexia, abnormal food intake and cachexia. The present
 sequence represents a nucleotide sequence from an example of the
 present invention.
 Sequence 83 BP: 29 A; 21 C; 8 G; 24 T;

Query Match 38.7%; Score 17.4; DB 1; Length 83;
 Best Local Similarity 64.9%; Pred. No. 3.7e+02;
 Matches 24; Conservative 1; Mismatches 12; Indels 0; Gaps 0

2 gacgcacatggcaactttaaaacacatggcagcat 38
 61 GAAGACTGTGGCAACTTACAGTAATTTGTAAATAAT 25

RESULT 12
 V16538
 ID V16538 standard; DNA; 86 BP.
 AC V16538;
 DT 11-JUN-1998 (first entry)
 DE Oligonucleotide V6635.
 KW Nucleic acid detection; Identification;
 KM variant diagnostic base sequence; ss.
 OS Synthetic.
 PN GB2312747-A.
 PD 05-NOV-1997.
 PF 29-APR-1997: 008581.
 PR 04-MAY-1996: GB-009441.
 PI (ZENNE) ZENNECA LTD.
 PI Brownie J, Little S, Whitcombe DM;
 PR WPI; 97-506147/47.
 PR Detection of a diagnostic base sequence in a nucleic acid sample -
 using primers with non-complementary tails, is useful for e.g.
 diagnosis of diseases and identification of samples in forensic
 medicine
 PS Disclosure; Page 24; 57pp; English.
 CC The present oligonucleotide was used to exemplify the method of the
 CC specification, which describes the detection of a diagnostic nucleic
 CC acid base. The method comprises contacting the sample with a diagnostic
 CC primer for the diagnostic base under hybridising conditions and in the
 CC presence of appropriate nucleoside triphosphates and an agent for
 CC polymerisation. The diagnostic primer has a non-complementary tail
 CC sequence comprising a tag and a detector region. Any extension products
 CC are used as a template for extension of a further primer which hybridises
 CC to a locus at a distance from the diagnostic base. The sample is
 CC contacted with a tag primer which selectively hybridises to the
 CC complement of the tag sequence in an extension product of the further
 CC primer. The presence or absence of the diagnostic base is detected by
 CC reference to the detector region in the further primer extension product.
 CC The method is used to identify one or more variant diagnostic base
 CC sequences against a background of normal diagnostic base sequences.
 CC Detection of nucleic acids is important for detection and diagnosis of
 CC alleles or DNA sequences associated or linked to genes responsible for
 CC genetic diseases in humans and other species; detection and diagnosis of
 CC neoplasms; detection and distinction between different pathogens,
 CC determining the purity of animal strains and pedigrees; and
 CC distinguishing and identifying different human and animal samples in
 CC forensic medicine.
 CC Sequence 86 BP: 22 A; 22 C; 25 G; 17 T;

```

Query Match          37.3%; Score 16.8; DB 1; Length 86;
Best Local Similarity 75.0%; Pred. No. 6.3e+02;
Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 2 gaccgatggcgaactttaaacacatat 29
    ||||| ||||| ||||| |||||
Db 53 GACGGCTTGGCACCATTTAAGAAATAT 80

RESULT 13
T88000/c
AC T88000 standard; DNA: 28 BP.
DE 27-APR-1998 (first entry)
DT primer used in isolation of Werner's syndrome 3 gene.
KW Werner's syndrome 3; WS-3; diagnosis; probe; examination;
KW prevention; antibody; reagent; human ontogenesis; detection;
KW PCR primer; ss.
OS Synthetic.
PN J09238683-A.
PD 16-SEP-1997.
PE 12-MAR-1996; 055144.
PR 12-MAR-1996; JP-055144.
PA (EITJ-) EJIN KENKUSHO KK.
DR WPI; 97-506552/47.
PT Werner's syndrome 3 gene - useful to develop products to detect.
PI diagnose and study the disease
PS Example 4: Page 22; 28pp; Japanese.
CC The present sequence was used in the isolation of the Werner's
CC syndrome 3 (WS-3) gene, which can be used as a diagnostic probe for
CC WS related diseases. It may also be used to examine and prevent
CC such diseases. The protein or a monoclonal or polyclonal antibody
CC raised against it can be used as a reagent for studies on human
CC ontogenesis, or WS-3 detection. An oligonucleotide probe specific
CC to the WS-3 gene can be used as a reagent for the detection of the
CC WS-3 gene.
SQ Sequence 28 BP; 7 A; 4 C; 8 G; 9 T;

Query Match          36.9%; Score 16.6; DB 1; Length 28;
Best Local Similarity 82.6%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 10 ggcacatttaaacacattgg 32
    | ||||| || ||| |||||
Db 28 GCCACTTCAACACATTATTGG 6

RESULT 14
V33999/c
AC V33999 standard; DNA: 47 BP.
AC V33999:
DE 25-JAN-1999 (first entry)
DT Forward primer for truncated apopain p12 subunit gene.
KW Inhibitor; apopain; binding site; crystal structure; primer; PCR;
KW amplification; Escherichia coli; apoptosis; ss.
OS Synthetic.
PN US5634228-A.
PD 10-NOV-1998.
PR 13-FEB-1997; 800007.
PE 13-FEB-1997; US-800007.
PA (MERI) MERCK & CO INC.
PI (MERI) MERCK FROST CANADA INC.
PI Becker JW, Fazil KM, Gallant M, Gareau Y, Labelle M,
PI Nicholson DM, Peterson EP, Rasper DM, Rotonda J,
PI Fuel R, Thornderry NA, Vaillancourt JP;
DR WPI; 99-008706/01.
PT Identifying inhibitors of apopain by rational drug design - from
PT ability to bind to the enzyme's substrate binding domain,
PT potentially useful for inhibiting apoptosis, e.g in treatment of
PT acquired immune deficiency syndrome, autoimmune disease, infections

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:22 ; Search time 4521.53 Seconds
(without alignments)
40.339 Million cell updates/sec

Title: US-09-164-714-15

Perfect score: 45

Sequence: 1 ggaagcagatgcaactttaa.....ctattggcagatcacgcttc 45

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *
44: gb_est25: *

45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_est48: *
83: gb_est49: *
84: gb_est50: *
85: gb_est51: *
86: em_est34: *
87: em_est35: *
88: em_est36: *
89: em_est37: *
90: gb_est52: *
91: gb_est53: *
92: gb_est54: *
93: gb_est55: *
94: gb_est56: *
95: em_est38: *
96: em_est39: *
97: em_est40: *
98: em_est41: *
99: em_est42: *
100: em_est43: *
101: em_est44: *
102: gb_est57: *
103: gb_est58: *
104: em_est45: *
105: gb_est59: *
106: gb_est60: *
107: gb_est61: *
108: gb_est62: *
109: gb_est63: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match Length	DB	ID	Description
1	19	42.2	82	41	AU006976
2	19	42.2	83	41	AU006976
3	18.6	41.3	87	42	AU006835
4	17.6	39.1	61	37	AU008855
5	17.6	39.1	92	23	R86574
6	17.4	38.7	85	20	D20939
7	17	37.8	97	42	A1130336
8	16.8	37.3	98	79	AM245015
9	16.6	36.9	66	23	R84752
10	16.6	36.9	70	42	AF090322
11	16.4	36.4	44	94	AQ025914
12	16.4	36.4	50	94	AA0207850
13	16.4	36.4	81	24	N28038
14	16.4	36.4	91	30	AA218354
15	16.4	36.4	97	35	AA591478
16	16.4	36.4	97	51	A1746501
17	16.4	36.4	100	29	A1152082
18	16.2	36.0	88	25	N90359
19	16.2	36.0	96	40	AA9680036
20	16.2	36.0	95	64	AM075131
21	16	35.6	71	64	B07636
22	16	35.6	73	64	AM080846
23	16	35.6	74	63	A1958931
24	16	35.6	94	74	AW201082
25	16	35.6	95	23	H48380
26	16	35.6	100	37	AA678059
27	15.8	35.1	40	46	A1462129
28	15.8	35.1	79	30	AA204291
29	15.8	35.1	85	83	AF020745
30	15.6	35.1	91	45	A1376009
31	15.6	34.7	79	34	AA515200
32	15.6	34.7	84	30	AA208974
33	15.6	34.7	85	39	AA876992
34	15.6	34.7	85	41	AU014419
35	15.6	34.7	85	41	AU014475
36	15.6	34.7	90	48	A1625619
37	15.6	34.7	93	48	H66572
38	15.6	34.7	97	61	A1818226
39	15.6	34.7	100	39	AA878683
40	15.6	34.7	100	62	A1936658
41	15.4	34.2	64	34	AA503991
42	15.4	34.2	81	40	DA3566
43	15.4	34.2	82	39	AA904068
44	15.4	34.2	87	83	AF020746
45	15.4	34.2	91	42	A1188513
					AU006976 AU006976
					AU006835 AU006835
					A1152611 uD89604.r.t
					AA700885 z140a01.s
					R86574 RABST149M
					D20939 HUMS01975
					A1130336 SMOVL3GCAN
					AM245015 BSBMFFS23
					R84752 yf28c09.r.t
					AF090323 AF0903232
					AQ025914 1(2)K0992
					AA207850 mw81e07.r
					N28038 EST000171.S
					AA218354 mw55c02.r
					AA591478 v156g02.r
					A1746501 u107g08.r
					A1152082 z148g12.r
					N90359 za15a04.r.t
					AA9680036 DDBS32.Peat
					AM075131 xa68f07.r
					B07636 CD1c8.Crl
					AM080846 XC38C09.X
					A1958931 tD23603.y
					AW201082 se97f05.y
					H48380 yf32b02.r.t
					AA678059 z125f07.s
					A1462129 uD70d04.x
					AA204291 nm32d08.r
					AF020745 Homo sapi
					A1376009 tE57e09.x
					AA515200 n154g07.s
					AA208974 nm72h12.r
					AA876992 z149f05.s
					AU014419 A0014475
					AU014475 A0014475
					A1625619 tV56h08.x
					H66572 yf72a06.r.t
					A

RESULT	1
AU006976	
LOCUS	AU006976 82 bp mRNA EST 31-JUL-1998
DEFINITION	AU006976 Schizosaccharomyces pombe late log phase cDNA
ACCESSION	Schizosaccharomyces pombe cDNA clone spc01229, mRNA sequence.
VERSION	AU006976
KEYWORDS	AU006976.1 GI:3343434
SOURCE	EST.
ORGANISM	fission yeast.
	Schizosaccharomyces pombe
	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
REFERENCE	Schizosaccharomycetaceae; Schizosaccharomyces.
AUTHORS	1 (bases 1 to 82)
TITLE	Morimyo,M. and Mita,K.
	Identification of expressed sequence tags of Schizosaccharomyces
JOURNAL	pombe
COMMENT	Unpublished (1998)
	On Nov 29, 1993 this sequence version replaced gi:636097.
	Contact: Mitsuo Morimyo
	Genome Research Group
	National Institute of Radiological Sciences

FEATURES		9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan	
SOURCE		Email: moriyo@nirs.go.jp	
1. .82		Location/Qualifiers	
/organism="Schizosaccharomyces pombe"			
/strain="972"			
/db_xref="taxon:4896"			
/clone="spc01229"			
/sex="h minus"			
/note="Vector: M13mp19; The cDNA library of Schizosaccharomyces pombe was prepared by cloning cDNA into the SmaI site of M13mp19 DNA and the direction of DNA sequences was not always from 5' to 3'. The cDNA data of Schizosaccharomyces pombe are available for searching on the World Wide Web. (URL, http://www.nirs.go.jp)"			
BASE COUNT		30 a 21 c 14 g 17 t	
ORIGIN			
Query Match		42.2%; Score 19; DB 41; Length 82;	
Best Local Similarity		71.4%; Pred. No. 3e+03;	
Matches		25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;	
Oy		5 gccatgcaactttaaacaactatggcagatc 39	
Db		46 gccatgcaacagcggaaaacgcctcatgsgaaaatacc 80	
RESULT 2			
AU006835		EST 31-JUL-1998	
LOCUS		AU006835 Schizosaccharomyces pombe late log phase cDNA	
DEFINITION		Schizosaccharomyces pombe cDNA clone spc01011, mRNA sequence.	
ACCESSION		AU006835	
VERSION		AU006835.1 GI:3343293	
KEYWORDS		EST.	
SOURCE		fission yeast.	
ORGANISM		Schizosaccharomyces pombe	
		Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;	
		Schizosaccharomycetaceae; Schizosaccharomyces.	
		1 (bases 1 to 83)	
REFERENCE		Moriyo,M. and Mita,K.	
AUTHORS		Identification of expressed sequence tags of Schizosaccharomyces pombe	
TITLE		Unpublished (1998)	
JOURNAL		On Jun 18, 1996 this sequence version replaced gi:136583.	
COMMENT		Contact: Mitsunori Moriyo	
		Genome Research Group	
		National Institute of Radiological Sciences	
		9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan	
		Email: moriyo@nirs.go.jp.	
FEATURES		Location/Qualifiers	
SOURCE		1. .83	
		/organism="Schizosaccharomyces pombe"	
		/strain="972"	
		/db_xref="taxon:4896"	
		/clone="spc01011"	
		/clone.lib="Schizosaccharomyces pombe late log phase cDNA"	
		/sex="h minus"	
		/note="Vector: M13mp19; The cDNA library of Schizosaccharomyces pombe was prepared by cloning cDNA into the SmaI site of M13mp19 DNA and the direction of DNA sequences was not always from 5' to 3'. The cDNA data of Schizosaccharomyces pombe are available for searching on the World Wide Web. (URL, http://www.nirs.go.jp)"	
BASE COUNT		30 a 21 c 14 g 17 t 1 others	
ORIGIN			
Query Match		42.2%; Score 19; DB 41; Length 83;	
Best Local Similarity		71.4%; Pred. No. 3e+03;	
Matches		25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;	

Qy	5	gccatgcaacttaaaacacatctggcagctatc	39
Db	46	gccattgcacacgaaaaacgctctatggaaatacc	80
RESULT	3		
LOCUS	A1152611/c		
DEFINITION	A1152611	97 bp	mRNA
ACCESSION	U089604.f1	Soares_NMPu	Mus musculus
VERSION	A1152611		CDNA clone IMAGE:1478046 5'
KEYWORDS			similar to TR:Q14731 Q14731 LAMININ AH ;, mRNA sequence.
SOURCE	A1152611.1	GI:3681080	
ORGANISM			EST.
REFERENCE			house mouse.
AUTHORS			Mus musculus
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
			Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
			1 (bases 1 to 97)
			Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
			Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
			Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
			Theising,B., Wille,T., Lennon,G., Soares,B., Wilson,R. and
			Waterston,R.
			The WashU-HMHI Mouse EST Project
			Unpublished (1996)
			On Apr 14, 1993 this sequence version replaced gi:639170.
COMMENT			Contact: Marra M/Mouse EST Project
			WashU-HMHI Mouse EST Project
			Washington University School of MedicineP
			4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
			Tel: 314 286 1800
			Fax: 314 286 1810
			Email: mouseest@watson.wustl.edu
			This clone is available royalty-free through LIND ; contact`the
			IMAGE Consortium (info@image.llnl.gov) for further information.
			MG1:926402
			Trace considered overall poor quality
			Possible reversed clone: similarity on wrong strand
			Seq primer: ~28ml3 rev2 ET from Amer sham
			High quality sequence stop: 1.
FEATURES			Location/Qualifiers
source			1..97
			/organism="Mus musculus"
			/db_xref="taxon:10090"
			/clone="IMAGE:1478046"
			/clone_id="Soares_NMPu"
			/sex="female"
			/dev_stage="adult"
			/lab_host="DH10B"
			/note="Organ: uterus; Vector: pF713D-Pac (Pharmacia) with
			a modified polylinker; 1st strand cDNA was prepared from
			pregnant mouse uterus, and was then primed with a Not I -
			oligo(9T) primer. Double-stranded cDNA was ligated to Eco
			RI adaptors (Pharmacia), digested with Not I and cloned
			into the Not I and Eco RI sites of the modified pF73
			vector. Library is normalized. Library was constructed by
			Bento Soares and M. Fatima Bonaldo."
BASE COUNT	25 a	19 c	21 g
ORIGIN			32 t
Query Match	41.3%	Score 18.6;	DB 42;
Best Local Similarity	65.9%	Pred. NO. 4.2e+03;	Length 97;
Matches	27;	Conservative 0;	Mismatches 14;
		Indels 0;	Gaps 0;
Qy	3	accgatgcaacttaaaacacatctggcagctatc	43
Db	77	aggcaatttcaattttaaccacatattgaaagactatct	37
RESULT	4		
			AA700885/c

LOCUS	AA700885	61 bp	mRNA	EST	19-DEC-1997
DEFINITION	c140a01.s1 Soares_fetal_liver_spleen_INFLS.S1 Homo sapiens cDNA clone IMAGE:452712 3', similar to SW:ID2.HUMAN Q02363 DNA-BINDING PROTEIN INHIBITOR ID-2. ; mRNA sequence.				
ACCESSION	AA700885				
VERSION	AA700885.1	GI:2704050			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 61) Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacey,M., Le,N., Lennon,G., Matra,M., Martin,J., Moore,B., Schellenberg,K., Stepec,M., Tan,F., Thaising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.				
TITLE	Washu-NCI human EST Project				
JOURNAL	Unpublished (1997)				
COMMENT	On Sep 12, 1996 this sequence version replaced gi:1397509. Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Trace considered overall poor quality Possible reversed clone: similarity on wrong strand Seq primer: -40m13 fwd. ET from Amersham High quality sequence stop: 1.				
FEATURES	Location/Qualifiers 1..61 /organism="Homo sapiens" /db_xref="GDB:1389068" /db_xref="taxon:9606" /clone IMAGE:452712 /clone_lib="Soares_fetal_liver_spleen_INFLS.S1" /sex="male" /dev_stage="20 week-post conception fetus" /lab_host="DH10B (ampicillin resistant)" /note="Organ: Liver and Spleen Vector: p7713D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen INFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACGGAAGATTAATTAAAGATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified p7713 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."				
BASE COUNT	g	a	g	c	21 g 22 t
ORIGIN					
Query Match	39.1%;	Score 17.6;	DB 37;	Length 61;	
Best Local Similarity	83.3%;	Pred. No.1e+04;	4;	Indels	0;
Matches	20;	Conservative	0;	Mismatches	0;
Oy	12	caacttaaaacactatgcag	35		
DB	56	CAACTTAACAACACTATTGTCAG	33		
RESULT	5				
LOCUS	R86574	92 bp	mRNA	EST	17-AUG-1995
DEFINITION	RABEN149W Rabbit Osteoclast, Dennis Sakai Oryctolagus cuniculus cDNA clone PABOC149 3', mRNA sequence.				
ACCESSION	R86574				
VERSION	R86574.1	GI:947228			
KEYWORDS	EST.				

SOURCE
ORGANISM Oryctolagus cuniculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1 (bases 1 to 92)
AUTHORS Sakai,D., Tong,H.-S. and Minkin,C.
TITLE Osteoclast Molecular Phenotyping by Random cDNA Sequencing
JOURNAL Bone 17 (2), 111-119 (1995)
MEDLINE 96021365
COMMENT On May 8, 1995 this sequence version replaced gi:800292.
Other ESTs: RABEST149T
Contact: Sakai D
Basic Sciences
University of Southern California
USC School of Dentistry, 925 West 34th Street, DEN-4220, Los
Angeles, CA 90089-0641
Tel: 2137405563
Fax: 2137407560
Email: sakai@molbio.usc.edu
Insert Length: 465 Std Error: 0.00
Seq primer: M13 forward
High quality sequence stop: 320.
Location/Qualifiers

FEATURES
source
1..92
/organism="Oryctolagus cuniculus"
/strain="New Zealand White"
/db_xref="taxon:9986"
/clone_lib="PRABOC149"
/lab_host="E. coli DH12S"
/note="Vector: pSPORT1; Site_1: Sali; Site_2: NotI;
Poly(A)+ RNA was purified from a 97% pure population of
osteoclasts prepared from the long bones of 10 day old
rabbits. First strand cDNA was synthesized by priming
with an oligo(dT)-NotI anchor-primer and second strand
cDNA was synthesized by replacement synthesis as described
by Gubler and Hoffman (Gene 25:283, 1983). Following the
addition of Sali adapters and NotI digestion, the cDNA was
cloned between the Sali (50) and NotI (30) sites of the
pSPORT1 (BRL) plasmid vector."

BASE COUNT
24 a 23 c 21 g 22 t 2 others

ORIGIN

Query Match 39.1%; Score 17.6; DB 23; Length 92;
Best Local Similarity 65.0%; Pred. No. 9.8e+03;
Matches 26; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Oy 3 acgcacatggcaacttaaaacacattggcagtcacgc 42
1 ||||| ||||| ||||| ||||| |||||
Db 3 AAGCCATGATACTCTTATTAACACGGTGGGATTCTCAAC 42

RESULT 6
D20993 85 bp mRNA EST 30-JUL-1996
LOCUS HUMS01975 Human promyelocyte Homo sapiens cDNA clone mp2228 3',
DEFINITION mRNA sequence.
ACCESSION D20993
VERSION D20993.1 GI:504813
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 85)
AUTHORS Okubo,K., Fukushima,A., Yoshii,J., Niiyama,T., Kojima,Y.,
Yoshinari,H., Arimoto,J., and Matsubara,K.
TITLE Gene expression of human promyelocytic cell line HL60 before and
after induction of differentiation. A new application of 3'directed
cDNA sequencing
JOURNAL Unpublished (1993)
COMMENT Contact: Okubo,K., Fukushima,A., Yoshii,J., Niiyama,T., Kojima,Y.,

Yoshinari,H., Arimoto,J. and Matsubara,K.
Institute for Molecular and Cellular Biology
Osaka University
3-1 Yamada-oka,Suita,Osaka 565,Japan.
location/Qualifiers

FEATURES
source
1..85
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="mp2228"
/clone_lib="Human promyelocyte"
/note="Female, adult, cell_line = HL60, cell_type =
promyelocyte."

BASE COUNT
27 a 18 c 15 g 22 t 3 others

ORIGIN

Query Match 38.7%; Score 17.4; DB 20; Length 85;
Best Local Similarity 68.6%; Pred. No. 1.2e+04;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Oy 1 ggacgcacatggcaacttaaaacacattggcagc 35
1 || ||| ||||| | ||||| |||
Db 11 GGAACACGCGACCTAGTATTCACCTATTAGCG 45

RESULT 7
A1130336 97 bp mRNA EST 26-MAR-1999
LOCUS A1130336
DEFINITION SMOWL3CAN17F11 Onchocerca volvulus infective larva cDNA
(SMA94WL-OVL3) Onchocerca volvulus cDNA clone onchl126 5' similar
to WP:044E4.4 CE08718 RNA-BINDING PROTEIN ; mRNA sequence.
ACCESSION A1130336
VERSION A1130336.1 GI:3600354
KEYWORDS EST.
SOURCE Onchocerca volvulus.
ORGANISM Onchocerca volvulus.
Eukaryota; Metazoa; Nematoda; Secernentea; Spirurida; Spirurida;
Filarioidea; Onchocercidae; Onchocerca.
REFERENCE 1 (bases 1 to 97)
AUTHORS Williams,S.A., Lizotte-Manlewski,M., Laney,S., Wenhong,L.,
Hillier,L., Allen,M., Bowles,L., Geisel,S., Jost,S., Kucaba,T.,
Martin,J., Steptoe,M., Theising,B., White,Y., Wylie,T.,
Chappell,J., Person,B., Gibbons,M., Harvey,N., Pape,D.,
Chamberlain,A., Morales,R., Schurk,R., Rittler,E., Kohn,S.,
Underwood,K. and Marra,M.
TITLE Molecular Parasitology OVL3
JOURNAL Unpublished (1998)
COMMENT On Oct. 30, 1996 this sequence version replaced gi:1656870.
Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith
College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genome@smith.edu
The library was constructed by Wenhong Lu. The library is available
from Dr. S.A. Williams, email genome@smith.edu when requesting this
clone from Dr. Williams, please reference the Williams lab clone id
- SMOVL3CAN17F11
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: 73 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES
source
1..97
/organism="Onchocerca volvulus"
/strain="Sierra Leone"
/db_xref="taxon:6282"
/clone_lib="onchl126"
/clone_lib="Onchocerca volvulus infective larva cDNA
(SMA94WL-OVL3)"
/lab_host="XL1-Blue MRF"

BASE COUNT 17 a 17 c 15 g 14 t 3 others
ORIGIN

and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonalido."

Query Match 36.9%; Score 16.6; DB 23; Length 66;
Best Local Similarity 59.5%; Pred. No. 2.3e+04;
Matches 25; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 ggaagcagatgcgaactttaaacactatgcagatcagc 42
||||| 11 11 11 11 11 11 11 11 11 11 11
Db 60 GGATGCGCAGTCNCACACATCAATTAATTTGGCAGTTGCGNC 19

RESULT 10

AF090232 70 bp mRNA EST 30-SEP-1998
LOCUS AF090232 Rat limonene-treated regressing tumor cDNA library Rattus

DEFINITION AF090232 Rat limonene-treated regressing tumor cDNA library Rattus
norvegicus cDNA clone mrg-14, mRNA sequence.

ACCESSION AF090232

VERSION AF090232.1 GI:3659534

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 70)

Identifying differential gene expression in monoterpane-treated

mammary carcinomas using subtractive display

J. Biol. Chem. 271 (46), 29286-29294 (1996)

97067187

On Sep 12, 1996 this sequence version replaced gi:1402086.

Contact: Ariazi, E.A. and Gould, M.N.

Oncology University of Wisconsin-Madison

600 Highland Ave., Room K4/3, Madison, WI 53792, USA

repressed by monoterpane.

FEATURES

Location/Qualifiers

1..70

/organism="Rattus norvegicus"

/strain="Wistar-Furth"

/db_xref="taxon:10116"

/clone="mrg-14"

/clone_lib="Rat limonene-treated regressing tumor cDNA

library"

/tissue_type="7,12-dimethylbenz[a]anthracene

(DMBA)-induced mammary carcinoma"

/note="limonene-treated regressing tumor cDNA library

subtracted from a control-treated non-regressing cDNA

library"

BASE COUNT 22 a 12 c 26 g 10 t

ORIGIN

Query Match 36.9%; Score 16.6; DB 42; Length 70;

Best Local Similarity 71.0%; Pred. No. 2.3e+04;

Matches 22; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ggaagcagatgcgaactttaaacactatgcagc 31

||||| 11 11 11 11 11 11 11 11 11 11 11

Db 37 GGAGGCGCAGTCGACGAGGAAGCAATTATG 67

RESULT 11

LOCUS A0025914

DEFINITION A0025914 44 bp DNA GSS 30-JUN-1998

melanogaster genomic Sequence recovered from Both 5' and 3' ends of

P element, genomic survey sequence.

ACCESSION A0025914

VERSION A0025914.1 GI:3266266

KEYWORDS GSS.

SOURCE fruit fly.

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 44)

Spreading, A.C., Stern, D., Beaton, A., Rehm, E.J., Laverly, T.,

Morden, N., Mista, S. and Rubin, G.M.

The BDP gene disruption project: Single P element insertions

mutating 30% of Drosophila autosomal genes

Unpublished (1998)

Contact: Gerald Rubin

Berkeley Drosophila Genome Project

University of California, Berkeley

LSA Building, Berkeley, CA 94720-3200, USA

Fax: 5106439947

Email: gerry@fruitfly.berkeley.edu

FEATURES

Location/Qualifiers

1..44

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone_lib="Drosophila melanogaster P lethal line"

/note="Inverse PCR was performed on Drosophila

melanogaster strains each of which contains a single P

transposable element insertion that is thought to cause

either lethality or sterility. The resultant fragment for

each strain was directly sequenced to determine the

genomic sequence at the site of insertion. Details of the

protocols used can be found at

http://fruitfly.berkeley.edu/P-disrupt/inverse_pcr.html."

BASE COUNT 11 a 11 c 11 g 10 t

ORIGIN

Query Match 36.4%; Score 16.4; DB 94; Length 44;

Best Local Similarity 94.4%; Pred. No. 2.8e+04;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 tggcaacttaaaacac 26

||||| 11 11 11 11 11 11 11 11 11 11 11

Db 14 TGGCACTTTATTAACAC 31

RESULT 12

LOCUS AA207850/C

DEFINITION AA207850 50 bp mRNA EST 12-MAR-1997

clone IMAGE:661476 5' similar to gb:62867 Mouse y box

transcription factor (MOUSE);, mRNA sequence.

ACCESSION AA207850

VERSION AA207850.1 GI:1804381

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 50)

Marrta, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Query Match	Best Local Similarity	Matches 26;	Conservative	0;	Mismatches	16;	Indels	0;	Gaps	0;
4	cgccatgccaactttaaaaacactattggcagctacagcttc	45								
1		11								
50	cgcatgcatgaatcaatcaagaacattttccagcaacatccaccttc	91								
BASE COUNT	27 a	25 c	17 g	22 t						
ORIGIN										

/sex="unknown"
 /tissue_type="fetus"
 /dev_stage="12.5dpc total fetus"
 /lab_host="VDH10B"
 /note="Organ: whole fetus; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGGCGCCCTTATTTTTTTTTTTTTTT 3'], on total mouse RNA [provided by Minoru Ko, Wayne State Univ.]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

```

Oligo dT. Gastrulating embryos were collected at 7.5dpfc
from C57Bl/6 x DBA matings, excluding embryos that had
developed head folds and all extraembryonic tissues.
Average insert size: 1.3 Kb (range: 0.5 - 3.0 Kb) .
Referenced in Development 121, 2479-2489 (1995)"

BASE COUNT      27 a      13 c      16 g      41 t
ORIGIN

Query Match      36.4%:  Score 16.4:  DB 35:  Length 97:
Best Local Similarity 76.9%:  Pred. No.2.7e+04:
Matches 20:  Conservative 0:  Mismatches 6:  Indels 0:  Gaps 0:

Oy      15  ctttaaacacactattgcagatca 40
      ||| ||||| ||| || ||
Db      96  CTTAAAAAACAGCTTGTGCATTATCA 71

Search completed: June 4, 2000, 13:53:25
Job time: 20949 sec

```

LOCUS	AA591478	97 bp	EST	16-SEP-1997
DEFINITION	v156902.r1 Beddington mouse embryonic region Mus musculus cDNA clone IMAGE:307826 5' similar to gb:M62867 Mouse y box transcription factor (MOUSE);, mRNA sequence.			
ACCESSION	AA591478			
VERSION	AA591478.1	GI:2404791		

SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE
AUTHORS
1 (bases 1 to 97)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Theising, B., Wylle, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE	The WashU-HHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	On Sep 12, 1996 this sequence version replaced g1:1393553.

WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through ILN; contact the
IMAGE Consortium (info@image.jnl.gov) for further information.
MGI:528490
Seq primer: -28m13 rev1 ET from Amersham.

FEATURES
source

```

/organism="Mus musculus"
/strain="C57BL6 x DBA"
/db_xref="taxon:10090"
/clone="IMAGE:907826"
/clone_lib="beddingington mouse embryonic region"
/sex="pooled"
/tissue_type="embryo"
/dev_stage="7.5dpc"
/lab_host="DH12S"
/note="Organ: whole embryo; Vector: pCMV-SPORT; Site:1;
Salt; Site:2: NotI; Cloned unidirectionally. Primer:

```

Oligo dT. Gastrulating embryos were collected at 7.5dpf from C57Bl6 x DBA matings, excluding embryos that had developed head folds and all extraembryonic tissues. Average insert size: 1.3 kb (range: 0.5 - 3.0 kb). Referenced in Development 121, 2479-2489 (1995).

BASE COUNT 27 a 13 c 16 g 41 t

ORIGIN

Query Match 36.4% Score 16.4: DB 35; Length 97;
Best Local Similarity 76.9% Pred. No. 2.7e+04;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps

Oy 15 ctttaaacactatgtgcagatracca 40
 ||| ||||| ||| ||| ||| |||
Db 96 ctttaaaacacgtttgtcatttaccaca 71

```

Oy      15  cttaaaaacactattggcagtatca 40
          ||| ||| ||| ||| ||| |||
Db      96  CTAAAAAACACGTTGTGCATTACCA 71

```

Search completed: June 4, 2000, 13:53:25
Job time: 20949 sec

1000 1000 1000 1000 1000 1000 1000 1000 1000 1000

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:45 ; Search time 244.64 Seconds
(without alignments)
23.910 Million cell updates/sec

Title: US-09-164-714-15

Perfect score: 45

Sequence: 1 ggaagccatgcaacttaaa.....ctatggcagtcacgttc 45

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/prodata/2/ina/5C.COMB.seq.*
4: /cgn2_6/prodata/2/ina/5D.COMB.seq.*
5: /cgn2_6/prodata/2/ina/6.COMB.seq.*
6: /cgn2_6/prodata/2/ina/PCTUS.COMB.seq.*
7: /cgn2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.4	40.9	45	6	PCT-US95-14639-6
2	17.8	39.6	47	3	US-08-039-198B-18
3	17.8	39.6	47	4	US-08-182-067-23
4	17.8	39.6	47	4	US-08-465-313-23
5	17.4	38.7	83	2	US-08-518-878B-47
6	17.4	38.7	83	2	US-08-294-522B-47
7	17.4	38.7	83	3	US-08-807-861A-47
8	17.4	38.7	83	3	US-08-470-868A-47
9	16.6	36.9	47	3	US-08-800-007A-3
10	16.6	36.9	47	3	US-08-800-007A-7
11	16.4	36.4	50	1	US-08-171-389-382
12	16.4	36.4	50	2	US-08-123-936-382
13	16.4	36.4	50	3	US-08-475-228A-382
14	16.4	36.4	50	5	US-08-482-080A-382
15	16.4	36.4	50	6	PCT-US93-12388-382
16	16.4	36.4	100	1	US-08-145-705A-2
17	16.2	36.0	83	2	US-08-518-878B-32
18	16.2	36.0	83	2	US-08-294-522B-32
19	16.2	36.0	83	3	US-08-807-861A-32
20	16.2	36.0	83	3	US-08-470-868A-32
21	15.8	35.1	33	4	US-08-242-098-6
22	15.8	35.1	32	2	US-08-600-999-13
23	15.8	35.1	37	2	US-08-616-398-2
24	15.6	34.7	51	3	US-08-636-113-1
25	15.4	34.2	41	5	US-09-109-063-45
26	15.4	34.2	44	5	US-09-109-063-44
27	15.4	34.2	45	6	PCT-US95-14639-12

28	15.4	34.2	50	1	US-08-171-389-503	Sequence 503, App
29	15.4	34.2	50	2	US-08-123-936-503	Sequence 503, App
30	15.4	34.2	50	3	US-08-475-228A-503	Sequence 503, App
31	15.4	34.2	50	5	US-08-482-080A-503	Sequence 503, App
32	15.4	34.2	50	6	PCT-US93-12388-503	Sequence 503, App
33	15.4	34.2	77	1	US-08-117-374-14	Sequence 14, Appl
34	15.4	34.2	77	1	US-08-280-263-14	Sequence 14, Appl
35	15.4	34.2	77	6	PCT-US94-10256-14	Sequence 14, Appl
36	15.4	34.2	87	1	US-08-105-483-403	Sequence 403, App
37	15.4	34.2	87	1	US-08-105-483-403	Sequence 403, App
38	15.4	34.2	87	2	US-08-709-209-404	Sequence 403, App
39	15.4	34.2	87	2	US-08-709-209-404	Sequence 404, App
40	15.4	34.2	87	2	US-08-303-275-107	Sequence 107, App
41	15.4	34.2	87	2	US-08-303-275-108	Sequence 108, App
42	15.4	34.2	87	2	US-08-458-101-403	Sequence 403, App
43	15.4	34.2	87	2	US-08-458-101-404	Sequence 404, App
44	15.4	34.2	87	3	US-08-417-210A-129	Sequence 129, App
45	15.4	34.2	87	3	US-08-417-210A-130	Sequence 130, App

ALIGNMENTS

RESULT 1

PCT-US95-14639-6

Sequence 6, Application PC/TUS9514639

GENERAL INFORMATION:

APPLICANT: PHILIPA MARRACK

APPLICANT: JOHN KAPLER

APPLICANT: RICHARD SHMONKEVITZ

APPLICANT: MASAZUMI MATSUMURA

TITLE OF INVENTION: PROTECTIVE EFFECTS OF MUTATED

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Swanson & Bratschun, L.L.C.

STREET: 8400 E. Prentice Ave., Suite 200

CITY: Englewood

STATE: Colorado

COUNTRY: USA

ZIP: 80111

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch

COMPUTER: IBM pc compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: WordPerfect 6.0 for windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/14639

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/338,373

FILING DATE: 14-November-1994

ATTORNEY/AGENT INFORMATION:

NAME: Barry J. Swanson

REGISTRATION NUMBER: 33,215

REFERENCE/DOCKET NUMBER: SUP020/PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 793-3333

TELEFAX: (303) 793-3433

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 45 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

PCT-US95-14639-6

Query Match

Best Local Similarity 78.6%; Score 18.4; DB 6; Length 45;

Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/182,067
FILING DATE: 23-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CB92/01289
FILING DATE: 15-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9115364.3
FILING DATE: 16-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BROOK, DAVID E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LYNX91-01A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: YES
US-08-465-313-23

Query Match 39.6%; Score 17.8; DB 4; Length 47;
Best Local Similarity 75.9%; Pred. No. 82;

Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 15 ctttaaacactatgacatgcagct 43
DB 11 ctttcataatgactgtgcagctgtaggt 39

RESULT 5
US-08-518-878B-47/c
Sequence 47, Application US/08518878B
Patent No. 5702902
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/518,878B
FILING DATE: 23-AUG-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-036
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:
LENGTH: 83 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-518-878B-47

Query Match 38.7%; Score 17.4; DB 2; Length 83;
Best Local Similarity 64.9%; Pred. No. 1.3e+02;
Matches 24; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

OY 2 gacgcacatgcaactttaaacactatgacgat 38
DB 61 GAAGATCTGGCACTTAGAAGTAAATTTGTAAAT 25

RESULT 6
US-08-294-522B-47/c
Sequence 47, Application US/08294522B
Patent No. 5741666
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: Compositions and Methods for the
Treatment of Body Weight Disorders, Including Obesity
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/294,522B
FILING DATE: 23-AUG-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 83 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-294-522B-47

Query Match 38.7%; Score 17.4; DB 2; Length 83;
Best Local Similarity 64.9%; Pred. No. 1.3e+02;
Matches 24; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

OY 2 gacgcacatgcaactttaaacactatgacgat 38
DB 61 GAAGATCTGGCACTTAGAAGTAAATTTGTAAAT 25

RESULT 7
US-08-807-861A-47/c
Sequence 47, Application US/08807861A
Patent No. 5853975

GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/807,861A
FILING DATE: 26-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/518,878
FILING DATE: 23-AUG-1995
APPLICATION NUMBER: US 08/470,868
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: US 08/294,522
FILING DATE: 23-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 83 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-807-861A-47

Query Match 38.7%; Score 17.4; DB 3; Length 83;
Best Local Similarity 64.9%; Pred. No. 1.3e+02;
Matches 24; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 2 gagcgcatggcaacttaaacactatggcgagat 38
|||:||||||||||| | ||| | ||
Db 61 GAAGAYCTGGCACTTTAGAGTAATTTGTAAAAAT 25

RESULT 8
US-08-470-868A-47/C
Sequence 47, Application US/08470868A
Patent No. 5861485
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis C.
TITLE OF INVENTION: Compositions and Methods for the
TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,868A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-0031-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 83 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-470-868A-47

Query Match 38.7%; Score 17.4; DB 3; Length 83;
Best Local Similarity 64.9%; Pred. No. 1.3e+02;
Matches 24; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 2 gagcgcatggcaacttaaacactatggcgagat 38
|||:||||||||||| | ||| | ||
Db 61 GAAGAYCTGGCACTTTAGAGTAATTTGTAAAAAT 25

RESULT 9
US-08-800-007A-3/C
Sequence 3, Application US/08800007A
Patent No. 5834228
GENERAL INFORMATION:
APPLICANT: Becker, Joseph
APPLICANT: Nicholson, Donald
APPLICANT: Rotonda, Jennifer
APPLICANT: Thornberry, Nancy
APPLICANT: Fazil, Kimberly
APPLICANT: Gallant, Michel
APPLICANT: Gareau, Yves
APPLICANT: Labelle, Marc
APPLICANT: Peterson, Erin
APPLICANT: Rasper, Dita
TITLE OF INVENTION: CRYSTAL STRUCTURE OF APOPAIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: JOSEPH A. COPPOLA - MERCK & CO., INC.
STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,007A
FILING DATE: 13-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: COPPOLA, JOSEPH A
REGISTRATION NUMBER: 38,413
REFERENCE/DOCKET NUMBER: 19644

TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-6734
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-800-007A-3

Query Match 36.9%; Score 16.6; DB 3; Length 47;
Best Local Similarity 71.0%; Pred. No. 2.4e+02;
Matches 22; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3 accgcagtgcaacttaaacacttggc 33
Db 44 ACCGCATGTCATCATCAACACCACTCATGAC 14

RESULT 10
US-08-800-007A-7/C
Sequence 7, Application US/08800007A
Patent No. 5834228
GENERAL INFORMATION:
APPLICANT: Becker, Joseph
APPLICANT: Nicholson, Donald
APPLICANT: Rotonda, Jennifer
APPLICANT: Thornderry, Nancy
APPLICANT: Fazil, Kimberly
APPLICANT: Gallant, Michel
APPLICANT: Gareau, Yves
APPLICANT: Labelle, Marc
APPLICANT: Peterson, Erin
APPLICANT: Rasper, Dita
TITLE OF INVENTION: CRYSTAL STRUCTURE OF APOPAIN
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: JOSEPH A. COPPOLA - MERCK & CO., INC.
STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,007A
FILING DATE: 13-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: COPPOLA, JOSEPH A
REGISTRATION NUMBER: 38,413
REFERENCE/DOCKET NUMBER: 19644
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-6734
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-800-007A-7

Query Match 36.9%; Score 16.6; DB 3; Length 47;
Best Local Similarity 71.0%; Pred. No. 2.4e+02;
Matches 22; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3 accgcagtgcaacttaaacacttggc 33
Db 44 ACCGCATGTCATCATCAACACCACTCATGAC 14

RESULT 11
US-08-171-389-382
Sequence 382, Application US/08171389
Patent No. 5578444
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 641
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/171,389
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175/G19P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 382:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human 70 kDa heat shock protein gene
US-08-171-389-382


```
Patent No. 6010849
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,080A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/171,389
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Brady, John F.
REGISTRATION NUMBER: 39,118
REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 382:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human 70 kDa heat shock protein gene
US-08-482-080A-382

Query Match          36.4%; Score 16.4; DB 5; Length 50;
Best Local Similarity 67.6%; Pred. No. 2.9e+02;
Matches 23; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```
OY 1 ggaagcattggcaacttaaaacactattggca 34
    ||| ||| ||| ||| ||| ||| ||| |||
Db 4 ggtctccgtgacgacttataaaacccaggggca 37
```

```
RESULT 15
PCT-US93-12388-382
; Sequence 382, Application PC/TUS9312388
; GENERAL INFORMATION:
```

```
APPLICANT:
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 641
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12388
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175.41/G19PCT2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 382:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human 70 kDa heat shock protein gene
PCT-US93-12388-382
```

```
Query Match          36.4%; Score 16.4; DB 6; Length 50;
Best Local Similarity 67.6%; Pred. No. 2.9e+02;
Matches 23; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```
OY 1 ggaagcattggcaacttaaaacactattggca 34
    ||| ||| ||| ||| ||| ||| ||| |||
Db 4 ggtctccgtgacgacttataaaacccaggggca 37
```

```
Search completed: June 4, 2000, 16:09:46
Job time: 28065 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:40 ; Search time 1236.38 Seconds

(without alignments)
-35.406 Million cell updates/sec

Title: US-09-164-714-16

Perfect score: 45
Sequence: 1 atcaagcttgatgagtga.....atgaagaagccaatgagcgc 45

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_om:*
5: gb_om:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pl2:*
11: gb_pl3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_un:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_om:*
23: em_om:*
24: em_om:*
25: em_om:*
26: em_om:*
27: em_om:*
28: em_om:*
29: em_om:*
30: em_om:*
31: em_om:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pl4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*

45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pl5:*
52: gb_pl8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	28.8	64.0	48	5	AR024453	AR024453 Sequence
2	25.6	56.9	45	5	A77178	A77178 Sequence 6
3	25.6	56.9	45	5	A77180	A77180 Sequence 6
4	25.6	56.9	45	5	I20213	I20213 Sequence 6
5	25.6	56.9	45	5	I20215	I20215 Sequence 8
6	25.4	56.4	31	5	A02249	A02249 Oligonucleo
7	25.4	56.4	32	5	AR014147	AR014147 Sequence
8	25.4	56.4	32	5	AR014350	AR014350 Sequence
9	25.4	56.4	42	5	A77174	A77174 Sequence 2
10	25.4	56.4	42	5	I20209	I20209 Sequence 2
11	25.4	56.4	54	5	A77184	A77184 Sequence 12
12	25.4	56.4	54	5	I20219	I20219 Sequence 12
13	24.4	54.2	32	5	AR051772	AR051772 Sequence
14	24.4	49.8	31	5	AR051765	AR051765 Sequence
15	22.4	49.8	39	5	AR005129	AR005129 Sequence
16	22.4	49.8	39	5	AR058256	AR058256 Sequence
17	22.4	49.8	50	5	A68390	A68390 Sequence 7
18	22.4	49.8	50	5	A68391	A68391 Sequence 8
19	22.4	49.8	61	10	S71567	S71567 catechol O-
20	22.4	49.8	63	5	AR038968	AR038968 Sequence
21	22.2	49.3	37	5	A77195	A77195 Sequence 23
22	22.2	49.3	37	5	I20230	I20230 Sequence 23
23	22.2	48.9	38	5	A79123	A79123 Sequence 7
24	22.2	48.9	38	5	A79124	A79124 Sequence 8
25	22.2	48.9	38	5	A92205	A92205 Sequence 7
26	22.2	48.9	38	5	A92206	A92206 Sequence 8
27	21.8	48.4	37	5	A77194	A77194 Sequence 22
28	21.8	48.4	37	5	I20229	I20229 Sequence 22
29	21.4	47.6	31	5	A02248	A02248 Oligonucleo
30	21.4	47.6	32	5	AR014146	AR014146 Sequence
31	21.4	47.6	32	5	AR014349	AR014349 Sequence
32	21.4	47.6	32	5	AR051764	AR051764 Sequence
33	21.4	47.6	48	5	I23381	I23381 Sequence 24
34	21.4	47.6	48	5	I33283	I33283 Sequence 24
35	21.4	47.6	48	5	A69104	A69104 Sequence 22
36	20.4	45.3	42	5	I23963	I23963 Sequence 12
37	20.4	45.3	42	5	I23964	I23964 Sequence 13
38	20.4	45.3	60	5	AR016577	AR016577 Sequence
39	20.4	45.3	63	5	I23344	I23344 Sequence 15
40	20.4	45.3	63	5	A49833	A49833 Sequence 6
41	20.2	44.9	48	5	A92221	A92221 Sequence 23
42	20.2	44.4	46	5	A69105	A69105 Sequence 23
43	19.8	44.0	27	5	AR026052	AR026052 Sequence
44	19.8	44.0	30	5	A37681	A37681 Sequence 52
45	19.8	44.0	30	5	AR036846	AR036846 Sequence

ALIGNMENTS

RESULT	1	AR024453	LOCUS	AR024453	48 bp	DNA	PAT	05-DEC-1998
DEFINITION		Sequence 15 from patent US 5795974.						
ACCESSION		AR024453						
VERSION		AR024453.1		GI:3977747				
KEYWORDS								
SOURCE		Unknown.						
ORGANISM		Unknown.						
REFERENCE		Unclassified.						
AUTHORS		1 (bases 1 to 48)						
TITLE		Cole,B.C., Atkin,C.L. and Knudtson,K.L.						
JOURNAL		Mycoplasma arthritis superantigen						
FEATURES		Patent: US 5795974-A 15 18-AUG-1998;						
source		Location/Qualifiers						
		1..48						
		"/organism="unknown"						
BASE COUNT		13 a		6 c		13 g		16 t
ORIGIN								

Query Match	64.0%	Score 28.8	DB 5	Length 48
Best Local Similarity	82.5%	Pred. No. 1.8		
Matches	33	Conservative	0	Mismatches 7; Indels 0; Gaps 0
Qy	4	aagcttagtagatggtgtagtgcgtaaaagccaatgagc	43	
Db	3	aagcttagtagatggtgtagtgcattatcttcaaaaagacc	42	

[illegible]

	Query Match	56.9%;	Score 25.6;	DB 5;	Length 45;
	Best Local Similarity	87.5%;	Pred. No. 25;		
Matches	28; Conservative	0;	Mismatches	4;	Indels 0; Gaps:
Oy	4 aagcttagtgcgtgcatgatgaataaacgcc 35 Db 3 AACCTTAGTCGATCGTGCATGTCC 34 				
RESULT	3				
LOCUS	A77180	45 bp	DNA	PAT	19-OCT-1999
DEFINITION	Sequence 8 from Patent EP0614989.				
ACCESSION	A77180				
VERSION	A77180.1	GI:6088887			
KEYWORDS					
SOURCE	.. unidentified. unidentified				
ORGANISM	unclassified.				

REFERENCE	1 (bases 1 to 45)
AUTHORS	Pleuecthun,A.D. and Kreibber,C.
TITLE	A METHOD FOR IN VIVO SELECTION OF LIGAND-BINDING PROTEINS
JOURNAL	Patent: EP 0614989-A 14-SEP-1994;
	MORPHOSYS PROTEINOPTIMERUNG (DE)
FEATURES	Location/Qualifiers
SOURCE	1..45
	/organism="unidentified"
	/db_xref="taxon:32644"
BASE COUNT	9 a 7 c 13 g 16 t
ORIGIN	
Query Match	56.9%; Score 25.6; DB 5; Length 45;
Best Local Similarity	87.5%; Pred. No. 25;
Matches	28; Conservative 0; Mismatches 4; Indels 0; Gaps 0
Oy	4 aaagctatgatgtgcatgtgcataaacgcc 35
db	3 AAGCTTAGTGATGATGCATGTGCATTTC 34

[illegible]

	Query Match	56.9%	Score 25.6;	DB 5;	Length 45;	
	Best Local Similarity	87.5%;	Pred. No. 25;			
	Matches 28;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;	
Oy	4 aaagctagtatgtagtgcgtgatgaataaacgc 35 					
Db	3 AAGCTTAGTGATGTGATGTGATGTTC 34 					
RESULT	5					
LOCUS	I20215					
DEFINITION	I20215 45 bp DNA					
ACCESSION	Sequence 8 from patent US 5514548.					
VERSION	I20215				PAT	07-OCT-1996
KEYWORDS	I20215.1 GI:1600570					
SOURCE	.					
ORGANISM	Unknown. Unknown.					
REFERENCE	Unclassified.					
AUTHORS	1 (bases 1 to 45)					
TITLE	Krebbber,K., Moroney,S., Pluckthu,n,A. and Schneider,C.					
JOURNAL	Method for in vivo selection of ligand-binding proteins					
FEATURES	Patent: US 5514548-A 8 07-MAY-1996; Location/Qualifiers					
source	1..45 /organism='unknown'					
BASE COUNT	9 a 7 c 13 g 16 t					
ORIGIN						

Query Match 56.9%; Score 25.6; DB 5; Length 45;
Best Local Similarity 87.5%; Pred. No. 25;
Matches 28; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 aagcttagtgatggtgatgtaaaagcc 35
|||||
Db 3 AAGCTTAGTGATGCTGATGATGATTTCC 34

RESULT 6
LOCUS A02249 31 bp DNA PAT 26-APR-1996
DEFINITION Oligonucleotide sequence (adaptor 9) from patent EP0282042.
ACCESSION A02249
VERSION A02249.1 GI:490303
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 31)
AUTHORS Doebl, H., Eggmann, B., Gentz, R., Hochuli, E. and Stueber, D.
TITLE Fusion proteins and their purification
JOURNAL Patent: EP 0282042-A 23 14-SEP-1988;
F. HOFMANN-LA ROCHE AG
FEATURES Location/Qualifiers
source 1..31
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 7 a 2 c 12 g 10 t
ORIGIN

Query Match 56.4%; Score 25.4; DB 5; Length 31;
Best Local Similarity 96.3%; Pred. No. 29;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 agcttagtgatggtgatgtaaa 31
|||||
Db 1 AGCTTAGTGATGCTGATGATGAGA 27

RESULT 7
LOCUS AR014147 32 bp DNA PAT 05-DEC-1998
DEFINITION Sequence 113 from patent US 5773249.
ACCESSION AR014147
VERSION AR014147.1 GI:3971601
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 32)
AUTHORS Cappello, J. and Ferrari, F.A.
TITLE High molecular weight collagen-like protein polymers
JOURNAL Patent: US 5773249-A 113 30-JUN-1998;
Location/Qualifiers
source 1..32
/organism="unknown"

BASE COUNT 9 a 2 c 12 g 9 t
ORIGIN

Query Match 56.4%; Score 25.4; DB 5; Length 32;
Best Local Similarity 96.3%; Pred. No. 29;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 agcttagtgatggtgatgtaaa 31
|||||
Db 1 AGCTTAGTGATGCTGATGATGAGA 27

RESULT 8
LOCUS AR014350

LOCUS AR014350 32 bp DNA PAT 05-DEC-1998
DEFINITION Sequence 25 from patent US 5773577.
ACCESSION AR014350
VERSION AR014350.1 GI:3971804
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 32)
AUTHORS Cappello, J.
TITLE Products comprising substrates capable of enzymatic cross-linking
JOURNAL Patent: US 5773577-A 25 30-JUN-1998;
Location/Qualifiers
FEATURES source 1..32
/organism="unknown"

BASE COUNT 9 a 2 c 12 g 9 t
ORIGIN

Query Match 56.4%; Score 25.4; DB 5; Length 32;
Best Local Similarity 96.3%; Pred. No. 29;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 agcttagtgatggtgatgtaaa 31
|||||
Db 1 AGCTTAGTGATGCTGATGATGAGA 27

RESULT 9
LOCUS A77174 42 bp DNA PAT 19-OCT-1999
DEFINITION Sequence 2 from Patent EP0614989.
ACCESSION A77174
VERSION A77174.1 GI:608881
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 42)
AUTHORS Plueckthun, A.D. and Krebber, C.
TITLE A METHOD FOR IN VIVO SELECTION OF LIGAND-BINDING PROTEINS
JOURNAL Patent: EP 0614989-A 14-SEP-1994;
MORPHOSYS PROTEINOPTIMIERUNG (DB)
FEATURES Location/Qualifiers
source 1..42
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 11 a 7 c 15 g 9 t
ORIGIN

Query Match 55.6%; Score 25; DB 5; Length 42;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 aagcttagtgatggtgatgta 28
|||||
Db 3 AAGCTTAGTGATGCTGATGATG 27

RESULT 10
LOCUS I20209 42 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 2 from patent US 5514548.
ACCESSION I20209
VERSION I20209.1 GI:1600564
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 42)
AUTHORS Krebber, K., Moroney, S., Plueckthun, A. and Schneider, C.
TITLE Method for in vivo selection of ligand-binding proteins

JOURNAL Patent: US 5514548-A 2 07-MAY-1996;
FEATURES Location/Qualifiers
source 1..42
/organism="unknown"
BASE COUNT 11 a 7 c 15 g 9 t
ORIGIN

Query Match 55.6%; Score 25; DB 5; Length 42;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 aagcttagtgatggtgatggtgatg 28
|||||
Db 3 AAGCTTAGTGATGCGTGTGATG 27

RESULT 11
A77184 54 bp DNA PAT 19-OCT-1999
LOCUS
DEFINITION Sequence 12 from Patent EP0614989.
ACCESSION A77184
VERSION A77184.1 GI:6088891
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 54)
AUTHORS Pluckthun,A.D. and Krebber,C.
TITLE A METHOD FOR IN VIVO SELECTION OF LIGAND-BINDING PROTEINS
JOURNAL Patent: EP 0614989-A 14-SEP-1994;
MORPHOSYS PROTEINOPTIMERUNG (DE)
FEATURES Location/Qualifiers
source 1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 14 a 6 c 21 g 13 t
ORIGIN

Query Match 55.6%; Score 25; DB 5; Length 54;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 aagcttagtgatggtgatggtgatg 28
|||||
Db 3 AAGCTTAGTGATGCGTGTGATG 27

RESULT 12
I20219 54 bp DNA PAT 07-OCT-1996
LOCUS
DEFINITION Sequence 12 from patent US 5514548.
ACCESSION I20219
VERSION I20219.1 GI:1600574
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 54)
AUTHORS Krebber,K., Moroney,S., Pluckthun,A. and Schneider,C.
TITLE Method for in vivo selection of ligand-binding proteins
JOURNAL Patent: US 5514548-A 12 07-MAY-1996;
FEATURES Location/Qualifiers
source 1..54
/organism="unknown"
BASE COUNT 14 a 6 c 21 g 13 t
ORIGIN

Query Match 55.6%; Score 25; DB 5; Length 54;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 aagcttagtgatggtgatggtgatg 28
|||||
Db 3 AAGCTTAGTGATGCGTGTGATG 27

RESULT 13
AR051772 32 bp DNA PAT 29-SEP-1999
LOCUS
DEFINITION Sequence 99 from patent US 5830713.
ACCESSION AR051772
VERSION AR051772.1 GI:5975136
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 32)
AUTHORS Ferrari,F.A., Cappello,J., Crissman,J.W. and Dorman,M.A.
TITLE Methods for preparing synthetic repetitive DNA
JOURNAL Patent: US 5830713-A 99 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..32
/organism="unknown"
BASE COUNT 8 a 2 c 12 g 10 t
ORIGIN

Query Match 54.2%; Score 24.4; DB 5; Length 32;
Best Local Similarity 96.2%; Pred. No. 67;
Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 gcttagtgatggtgatggtgatgaa 31
|||||
Db 2 GCTTAGTGATGCGTGTGATGAGA 27

RESULT 14
AR051765 31 bp DNA PAT 29-SEP-1999
LOCUS
DEFINITION Sequence 83 from patent US 5830713.
ACCESSION AR051765
VERSION AR051765.1 GI:5975129
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 31)
AUTHORS Ferrari,F.A., Cappello,J., Crissman,J.W. and Dorman,M.A.
TITLE Methods for preparing synthetic repetitive DNA
JOURNAL Patent: US 5830713-A 83 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..31
/organism="unknown"
BASE COUNT 9 a 1 c 12 g 9 t
ORIGIN

Query Match 49.8%; Score 22.4; DB 5; Length 31;
Best Local Similarity 95.8%; Pred. No. 3.5e+02;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 ttagtgatggtgatggtgatgaa 31
|||||
Db 3 TTAGTGATGCGTGTGATGAGA 26

RESULT 15
AR005129 39 bp DNA PAT 04-DEC-1998
LOCUS
DEFINITION Sequence 34 from patent US 5747452.
ACCESSION AR005129
VERSION AR005129.1 GI:3966008
KEYWORDS

Unclassified.

III peptides

Source	1. .39
1. .39	

8 ttatnatatnatatnaaa 31

Search completed: June 4, 2000, 16:05:41
Job time: 27893 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:25 ; Search time 4521.53 Seconds
(without alignments)
40.339 Million cell updates/sec

Title: US-09-164-714-16

Perfect score: 45
Sequence: 1 atcaagcttgatggtgta.....atgaagccaatgagcgc 45

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202661650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 08
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_est48:*
83: gb_est49:*
84: gb_est50:*
85: gb_est51:*
86: em_est35:*
87: em_est36:*
88: em_est37:*
89: em_est38:*
90: gb_est52:*
91: gb_est53:*
92: gb_est54:*
93: gb_est55:*
94: gb_est56:*
95: em_est39:*
96: em_est40:*
97: em_est41:*
98: em_est42:*
99: em_est43:*
100: em_est44:*
101: em_est45:*
102: gb_est57:*
103: gb_est58:*
104: em_est46:*
105: gb_est59:*
106: gb_est60:*
107: gb_est61:*
108: gb_est62:*
109: gb_est63:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
C 1	20	44.4	100	94	A0026191	A0026191 1(3)03847
C 2	18.8	40.4	76	44	A1313567	A1313567 EST021 Mo
C 3	18.2	40.4	90	51	A0025298	A0025298 EP(3)3123
C 4	18.2	40.4	90	69	A1748511	A1748511 sb5a04.y
C 5	18.2	40.4	91	45	AM119559	AM119559 sda8a04.y
C 6	18.2	40.4	91	45	A1366782	A1366782 qqa4d10.x
C 7	18.2	40.4	100	39	AA906127	AA906127 c93c04.s
C 8	18.2	39.6	83	28	AA108264	AA108264 EST0007 r
C 9	17.6	39.1	73	50	A1708585	A1708585 as98g06.x
C 10	17.6	39.1	84	27	AA038117	AA038117 m181g04.r
C 11	17.4	38.7	58	36	AA663265	AA663265 ab80f10.s
C 12	17.4	38.7	87	39	AA861714	AA861714 ak30h09.s
C 13	17.4	38.7	88	47	A1522508	A1522508 fb20d05.x
C 14	17.2	38.2	89	23	L46917	L46917 SCMRAP004 G
C 15	17.2	38.2	100	63	A1963567	A1963567 wr64h11.x
C 16	17	37.8	84	83	AF039763	AF039763 Homo sapi
C 17	17	37.8	100	35	AA537389	AA537389 vj98a07.r
C 18	16.8	37.3	65	41	AU008149	AU008149 AU008149
C 19	16.8	37.3	65	41	AU008150	AU008150 AU008150
C 20	16.8	37.3	88	39	AA885639	AA885639 c932b08.s
C 21	16.8	37.3	94	46	A1460411	A1460411 sab1c02.y
C 22	16.8	37.3	100	39	AA902339	AA902339 ok92c10.s
C 23	16.8	37.3	100	62	A1921625	A1921625 wo27h05.x
C 24	16.6	36.9	40	45	AJ237265	AJ237265 AJ237265
C 25	16.6	36.9	59	27	AA026247	AA026247 zj99a04.s
C 26	16.6	36.9	63	82	GA246098	GA246098 Gallus ga
C 27	16.6	36.9	71	44	A1310271	A1310271 tba47h04.x
C 28	16.6	36.9	82	68	HSM002994	HSM002994 Homo sapi
C 29	16.6	36.9	86	35	AA574367	AA574367 nt46b08.s
C 30	16.6	36.9	86	82	CNS00W22	AL0944220 Arabidops
C 31	16.6	36.9	88	68	HSM002995	AL038519 Homo sapi
C 32	16.4	36.4	61	33	AA417038	AA417038 znl3a02.s
C 33	16.4	36.4	85	26	W99173	W99173 ME91604.r1
C 34	16.4	36.4	85	33	AA443226	AA443226 aal1a08.r
C 35	16.4	36.4	90	50	AU060644	AU060644 AU060644
C 36	16.4	36.4	91	21	T96074	T96074 ye47a12.s1
C 37	16.4	36.4	96	84	B34289	B34289 HS-1024-A2-
C 38	16.2	36.0	46	37	AA714376	AA714376 nw20b06.s
C 39	16.2	36.0	65	81	AA423329	AA423329 sh66c03.y
C 40	16.2	36.0	67	63	A1037391	A1037391 ub52b11.r
C 41	16.2	36.0	67	63	A1988376	A1988376 sdt01b03.y
C 42	16.2	36.0	70	34	AA467126	AA467126 vd98d01.r
C 43	16.2	36.0	71	45	A1348869	A1348869 tdb05b07.x
C 44	16.2	36.0	73	33	AA445034	AA445034 vg57g06.r
C 45	16.2	36.0	73	34	AA518487	AA518487 v101b09.r

ALIGNMENTS

RESULT 1
LOCUS A0026191/c 100 bp DNA 30-JUN-1998
DEFINITION I(3)03847 Drosophila melanogaster P lethal line Drosophila
melanogaster genomic Sequence recovered from Both 5' and 3' ends of
P element, genomic survey sequence.

ACCESSION A0026191
VERSION A0026191.1 GI:3266476
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 100)
AUTHORS Spradling,A.C., Stern,D., Beaton,A., Rehm,E.J., Laverly,T.,
Mozden,N., Mistra,S. and Rubin,G.M.
TITLE The BDP gene disruption project: Single P element insertions
mutating 30% of Drosophila autosomal genes
JOURNAL Unpublished (1998)
COMMENT Contact: Gerald Rubin

Berkeley Drosophila Genome Project
University of California, Berkeley
LSA Building, Berkeley, CA 94720-3200, USA
Fax: 5106439947
Email: gerryl@fruitfly.berkeley.edu

Sequence recovery method was inverse PCR.

Sequence orientation is forward strand relative to 5' end of P
element

The P element insertion position is base 052 in the 100 bases. This
insertion position refers to the first base of the 8 base target
recognition sequence.

Class: transposon-tagged.

FEATURES

source

Location/Qualifiers

1..100
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

/clone_lib="Drosophila melanogaster P lethal line"
/note="Inverse PCR was performed on Drosophila

melanogaster strains each of which contains a single P
transposable element insertion that is thought to cause
either lethality or sterility. The resultant fragment for
each strain was directly sequenced to determine the
genomic sequence at the site of insertion. Details of the
protocols used can be found at
http://fruitfly.berkeley.edu/P-disrupt/Inverse_pcr.html."

BASE COUNT
ORIGIN

20 a 35 c 21 g 24 t

Query Match 44.4% Score 20; DB 94; Length 100;

Best Local Similarity 65.9%; Pred No. 1.6e+03;

Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 2 tcaagcttagtgatgtagtgaatgaacccaatgagc 45
||||| ||||| || || ||||| || |||||
DB 44 tcaagatagtgatgtagaggaatgaatgctgacgc 1

RESULT 2
LOCUS A1313567/c 68 bp mRNA EST 15-DEC-1998
DEFINITION EST021 Mouse ES cell lambda zap Express Library Mus musculus cDNA
clone HBMG015R 3', mRNA sequence.
ACCESSION A1313567
VERSION A1313567.1 GI:4027734
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 68)
AUTHORS Kim,S.J., Shin,J.H., Kim,J.H., Kim,S.H., Chae,J.H., Park,E.J.,

Hong,S.H., Park,S.D., Jeong,S.J. and Kim,C.G.
Isolation of Developmentally Regulated Novel Genes Based on

Sequence Identity and Gene Expression Pattern
Unpublished (1998)

On Aug 26, 1998 this sequence version replaced gi:3475005.
COMMENT Contact: Chul Geun Kim

Molecular Genetics
University of Hanyang, Department of Life Science

Haengdang dong 17, Sungdong gu, Seoul 133-791, Korea (ROK)
Tel: 82-2-2290-0957

Fax: 82-2-2296-5996
Email: cgm@emall.hanyang.ac.kr

homology chosen cDNA clone from mouse ES cell cDNA library with
Seq primer: 73 Reverse
High quality sequence stop: 68.

FEATURES

Location/Qualifiers

```

source
1. .68
/organism="Mus musculus"
/strain="129/sv"
/db_xref="taxon:10090"
/clone_lib="HBMG015R"
/clone_id="Mouse ES cell Lambda Zap Express Library"
/sex="male"
/tissue_type="blastocyst inner cell mass"
/cell_type="embryonic stem cell"
/dev_stage="blastocyst"
/note="Vector: Uni-Zap XR Phage and generate
pbuescript-SK phagemid vector by in vivo excision."
BASE COUNT
19 a 18 c 13 g 18 t
ORIGIN

Query Match
Best Local Similarity 41.8%; Score 18.8; DB 44; Length 68;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 atcaagcttagtgatggtgatgataa 30
1 ||| ||| ||| ||| ||| ||| |||
50 ACCAAGCTTGAGATGAAGATGGTGGAA 21

RESULT 3
A0025298 76 bp DNA GSS 14-OCT-1998
LOCUS
DEFINITION
EP(3)3123-Sprime Drosophila melanogaster EP line Drosophila
melanogaster genomic Sequence recovered from 5' end of P element,
genomic survey sequence.
ACCESSION
A0025298
VERSION
A0025298.1 GI:3265650
KEYWORDS
GSS.
SOURCE
fruit fly.
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 76)
Rehm,E.J. and Rubin,G.M.
The BDGP gene disruption project: single EP element insertions
Unpublished (1998)
Contact: Gerald Rubin
Berkeley Drosophila Genome Project
University of California, Berkeley
LSA Building, Berkeley, CA 94720-3200, USA
Fax: 5106439947
Email: gerry@fruitfly.berkeley.edu

Sequence recovery method was inverse PCR.

element
Sequence orientation is forward strand relative to 5' end of P
element

The P element insertion position is base 69 in the 76 bases. This
insertion position refers to the first base of the 8 base target
recognition sequence.

Class: transposon-tagged.
location/Qualifiers
1. .76
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="Drosophila melanogaster EP line"
/note="Inverse PCR was performed on Drosophila
melanogaster strains each of which contains a single EP
transposable element insertion. (The generation of these
insertion strains is described in North P, Szabo K, Bailey
A, Laverly T, Rehm J, Rubin GM, Weigmann K, Millan M, Benes
V, Ansoorge W, Cohen SM. 1998. Systematic gain-of-function
genetics in Drosophila. Development 6:1049-1057.) The
resultant fragment for each strain was directly sequenced

```

```

to determine the genomic sequence at the site of
insertion. Details of the protocols used can be found at
http://fruitfly.berkeley.edu/P_disrupt/inverse_per.html."
BASE COUNT
18 a 14 c 19 g 25 t
ORIGIN

Query Match
Best Local Similarity 40.4%; Score 18.2; DB 94; Length 76;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 8 ttatgatggtgatgataaagccaa 38
1 ||| ||| ||| ||| ||| ||| |||
45 TTGTGCTTGTGATGTGAATGAAGCCAA 75

RESULT 4
A1748511/c 90 bp mRNA EST 13-DEC-1999
LOCUS
DEFINITION
Gm-c1016-295 5', mRNA sequence.
ACCESSION
A1748511
VERSION
A1748511.1 GI:5126775
KEYWORDS
EST.
SOURCE
soybean.
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoidae;
Glycine.
1 (bases 1 to 90)
Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Coryell,V.,
Khanina,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N.,
Schurr,R., Ritter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterson,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
On Jun 5, 1998 this sequence version replaced gi:3188332.
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com.
location/Qualifiers
1. .90
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1016-295"
/clone_lib="Gm-c1016"
/tissue_type="Immature flowers of field grown plants"
/lab_host="X110-Gold"
/note="Vector: pBluescript II XR. Site_1: EcoRI. Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from immature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pBluescript II
XR library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pBluescript vector. The ligated cDNA fragments were
transformed into X110-Gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John

```

[illegible]

QUERY	Query Match Best Local Similarity Matches	40.4%; 66.7%; 26;	Score 18.2; Pred. No. 6.5e+03; 0;	DB 69; Mismatched 13;	Length 90; Indels 0;
Db	76	ACGCTTGTGGTTGATTAGTTCAGAAATATCCAAAACAG	38		
OY	4	aagcttagtcatgtgatgcatgaagaacccaatgag	42		
RESULT	6				
LOCUS	AI366782	91 bp	mRNA	EST	13-FEB-1999
DEFINITION	gg41d10.x1 Soares_NHMPU_S1 Homo sapiens			cDNA clone IMAGE:1935091	
ACCESSION	AI366782				
VERSION	AI366782.1	GI:4136527			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
	Eutheraia; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1 (bases 1 to 91)				
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	On Jan 14, 1998 this sequence version replaced gi:1798622. Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Insert Length: 725 Std Error: 0.00 Seq primer: -40UP from Glbpco High quality sequence stop: 71. Location/Qualifiers				
FEATURES	source	1..91			
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:1935091"				
	/clone_id="Soares_NHMPU_S1"				
	/tisue_type="Soares_NHMPU_S1"				
	/tisue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"				
	/lab_host="DH10B"				
	/note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbM, pregnant uterus NbHPU, and fetal heart NBH129W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."				
BASE COUNT	35 a	10 c	17 g	29 t	
ORIGIN					
Query Match	40.4%;	Score 18.2;	DB 69;	Length 91;	
Best Local Similarity	66.7%;	Pred. No. 6.6e+03;			
Matches	26;	Conservative 0;	Mismatched 13;	Indels 0;	Gaps 0;
OY	3	caagcttagtcatgtgatgcatgaagaacccaatga	41		
Db	51	CAGCATGTATGTGGAGAAGGTTAAGTCAACCACACATGA	89		
RESULT	7				

[illegible]

TITLE	Rebillard,G., Vago,P., Pujoil,R. and Hamel,C.P.
JOURNAL	Identification of preferentially expressed cochlear genes by systematic sequencing of a rat cochlea cDNA library
MEDLINE	Brain Res. Mol. Brain Res. 47 (1-2), 1-10 (1997)
COMMENT	On Sep 12, 1996 this sequence version replaced gi:1393651. Contact: Hamel, CP Laboratoire de Neurobiologie de l'Audition - INSERM U254 Institut National de la Sante et de la Recherche Medicale C.H.R. St-Charles, 34295 Montpellier cedex, France Tel: (33) 04 67 33 69 75 Fax: (33) 04 67 52 56 01 Email: biomol@nmet.fr
FEATURES	PCR Primers FORWARD: pUC/M13 reverse primer BACKWARD: pUC/M13 forward primer Seq primer: pUC/M13 reverse primer High quality sequence stop: 83. Location/Qualifiers 1..83 /organism="Rattus norvegicus" /strain="Wistar" /db_xref="taxon:10116" /clone="PCO38" /clone_lib="rat lambda ZAPII library (C.P.Hamel)" /dev_stage="postnatal day 24" /lab_host="Escherichia coli" /note="Organ: cochlea; Vector: lambda ZAPII; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT	13 a 34 c 17 g 18 t 1 others
ORIGIN	
Query Match	39.6%; Score 17.8; DB 28; Length 83;
Best Local Similarity	67.6%; Pred. No. 8.9e+03;
Matches 25; Conservative	0; Mismatches 12; Indels 0; Gaps 0;
Dd 7	cctagtgatgltgatgtcatgaagaacaaatgagc 43
A1708585	45 CATGGTGTGGCGATGATGAGAAAGACCTCGAGGGGC 9
LOCUS	73 bp mRNA EST 04-JUN-1999
DEFINITION	as98g906.x1 Barstead aorta HPLB6 Homo sapiens CDNA clone
REGION (HUMAN);	IMAGE:2353690 3' similar to gb:L03555 IG KAPPA CHAIN PRECURSOR V-I mRNA sequence.
ACCESSION	A1708585
VERSION	A1708585.1 GI:4998361
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 73) Hillier,L., Allen,M., Bowles,L., Duboue,T., Giesel,G., Jost,S., Kizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Stepien,M., Tan,F., Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R. MashU-NCI human EST Project Unpublished (1997) On Dec 20, 1995 this sequence version replaced gi:1133755. Contact: Wilson RK Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -40UP from Gluco High quality sequence stop:1.
TITLE	
JOURNAL	
COMMENT	

```
FEATURES
SOURCE
    Location/Qualifiers
    1..73
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:2353690"
    /clone_lib="Barstead aorta HPLRB6"
    /sex="male"
    /dev_stage="adult, age 64"
    /lab_host="DH10B (phage resistant)"
    /note="Organ: aorta; Vector: pRT3D-Pac (Pharmacia) with a
    modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
    strand cDNA was primed with a Not I - oligo(dT) primer (5'
    TGTTACGATCTGAAGTGGAGCGGCCGCTTTTCTTTTCTTTTCTTTT
    3'); double-stranded cDNA was ligated to Eco RI adaptors
    [5', AATTCGATCGAC 3' and 5', GTTGCATCGG 3'], digested
    with Not I and cloned into the Not I and Eco RI sites of
    the modified pRT3 vector. Library constructed by Bob
    Barstead."
BASE COUNT      19 a      13 c      16 g      21 t      4 others
ORIGIN

Query Match      39.1%; Score 17.6; DB 50; Length 73;
Best Local Similarity 67.6%; Pred.No. 1e+04;
Matches 23; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 atcaagcttagtgatgtatgtatgaagaac 34
      ||||| ||| | ||||| ||| | |||
DB      35 ATCANGCTGCANGCTGCTGATGTGTAAGTATC 68

RESULT 10
LOCUS      AA038117      84 bp      mRNA      EST      28-AUG-1996
DEFINITION mi81g04.r1 Soares mouse p3MMP19.5 Mus musculus cDNA clone
            IMAGE:473046.5' similar to gb:V00714 Mouse gene for alpha-globin
            (MOUSE);, mRNA sequence.
ACCESSION      AA038117
VERSION      AA038117.1 GI:1513526
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 84)
AUTHORS      Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
            The WashU-HMNI Mouse EST Project
            Unpublished (1996)
            On Sep 21, 1992 this sequence version replaced gi:279433.
            Contact: Marra M/Mouse EST Project
            WashU-HMNI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@wustl.edu
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:283790
            Trace considered overall poor quality
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 1.
            Location/Qualifiers
            1..84
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /clone="IMAGE:473046"
            /clone_lib="Soares mouse p3MMP19.5"
            /dev_stage="19.5 dpc total fetus"

FEATURES
SOURCE
    Location/Qualifiers
    1..73
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:2353690"
    /clone_lib="Barstead aorta HPLRB6"
    /sex="male"
    /dev_stage="adult, age 64"
    /lab_host="DH10B (phage resistant)"
    /note="Organ: aorta; Vector: pRT3D-Pac (Pharmacia) with a
    modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
    strand cDNA was primed with a Not I - oligo(dT) primer (5'
    TGTTACGATCTGAAGTGGAGCGGCCGCTTTTCTTTTCTTTTCTTTT
    3'); double-stranded cDNA was ligated to Eco RI adaptors
    [5', AATTCGATCGAC 3' and 5', GTTGCATCGG 3'], digested
    with Not I and Eco RI sites of the modified pRT3 vector.
    Library constructed by Bob Barstead."
BASE COUNT      19 a      13 c      16 g      21 t      4 others
ORIGIN

Query Match      39.1%; Score 17.6; DB 27; Length 84;
Best Local Similarity 71.9%; Pred.No. 1e+04;
Matches 23; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      10 agtgatggtgatgtatgaagaacccaatga 41
      ||||| ||| | ||||| ||| | |||
DB      29 ATTGGTGGCTATGGTATGATGATATGGAGATGA 60

RESULT 11
LOCUS      AA663265      58 bp      mRNA      EST      12-NOV-1997
DEFINITION ab06f10.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
            IMAGE:853291.3' similar to SW:PR8_MOUSE P49014 265 PROTEASE
            REGULATORY SUBUNIT 4 ;, mRNA sequence.
ACCESSION      AA663265
VERSION      AA663265.1 GI:2617256
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 58)
AUTHORS      Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisler,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theisinger,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
            WashU-NCI human EST Project
            Unpublished (1997)
            On Sep 12, 1996 this sequence version replaced gi:1394008.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@wustl.edu
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
            Location/Qualifiers
            1..58
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:853291"
            /clone_lib="Stratagene fetal retina 937202"
            /sex="mixed"
            /lab_host="SOLR (kanamycin resistant)"
            /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
            XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled
            retinal tissue. Average insert size: 1.0 kb. Uni-ZAP XR
            Vector; -5' adaptor sequence: 5' GATTCGGCAGAG 3' -3'
            adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'."

BASE COUNT      16 a      13 c      17 g      12 t
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:46 ; Search time 244.64 Seconds
(without alignments)
23.910 Million cell updates/sec

Title: US-09-164-714-16

Perfect score: 45

Sequence: 1 atcaagcttgatgtgtga.....atgaagaagccaatgagcgc 45

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/6.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/PCUS.COMB.seq:*
7: /cgn2_6/ptodata/2/1na/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28.8	64.0	48	2	US-08-621-081A-15
2	25.6	56.9	45	1	US-08-197-770-6
3	25.6	56.9	45	1	US-08-197-770-8
4	25.4	56.4	32	2	US-08-642-255-113
5	25.4	56.4	32	2	US-08-397-633A-25
6	25.5	55.6	42	1	US-08-197-770-2
7	25.5	55.6	42	1	US-08-197-770-12
8	24.6	54.7	64	3	US-08-697-631A-12
9	24.6	54.7	65	3	US-08-697-631A-13
10	24.6	54.7	66	3	US-08-697-631A-11
11	24.6	54.7	66	3	US-08-697-631A-14
12	24.4	54.2	32	3	US-08-707-237A-99
13	24.4	54.2	32	3	US-08-642-246-21
14	24.4	54.2	32	6	PCT-US96-06229-21
15	22.4	49.8	31	2	US-08-707-237A-83
16	22.4	49.8	31	2	US-08-459-064B-34
17	22.4	49.8	39	3	US-08-460-421A-34
18	22.4	49.8	63	3	US-08-416-336-3
19	22.2	49.3	37	1	US-08-197-770-23
20	21.8	48.4	37	1	US-08-197-770-22
21	21.4	47.6	32	2	US-08-642-255-112
22	21.4	47.6	32	2	US-08-397-633A-24
23	21.4	47.6	32	2	US-08-707-237A-82
24	21.4	47.6	32	5	US-08-642-246-20
25	21.4	47.6	32	6	PCT-US96-06229-20
26	21.4	47.6	54	5	US-08-989-370-10
27	21.4	46.7	48	1	US-08-221-817-24

28	21	46.7	48	1	US-08-454-439-24	Sequence 24, Appl
29	21	46.7	48	6	PCT-US94-10487-24	Sequence 24, Appl
30	21	46.7	54	5	US-08-989-370-11	Sequence 11, Appl
31	20.6	45.8	54	5	US-08-989-370-7	Sequence 7, Appl
32	20.4	45.3	42	1	US-08-250-859-12	Sequence 12, Appl
33	20.4	45.3	42	1	US-08-250-859-13	Sequence 13, Appl
34	20.4	45.3	42	2	US-08-490-803-12	Sequence 12, Appl
35	20.4	45.3	42	2	US-08-490-803-13	Sequence 13, Appl
36	20.4	45.3	42	4	US-08-457-254-23	Sequence 23, Appl
37	20.4	45.3	42	4	US-08-457-254-24	Sequence 24, Appl
38	20.4	45.3	42	4	US-08-484-257-12	Sequence 12, Appl
39	20.4	45.3	42	4	US-08-484-257-13	Sequence 13, Appl
40	20.4	45.3	42	6	PCT-US94-08806-12	Sequence 12, Appl
41	20.4	45.3	42	6	PCT-US94-08806-13	Sequence 13, Appl
42	20.4	45.3	42	6	PCT-US95-01775-12	Sequence 12, Appl
43	20.4	45.3	42	6	PCT-US95-01775-13	Sequence 13, Appl
44	20.4	45.3	42	6	PCT-US95-16626-23	Sequence 23, Appl
45	20.4	45.3	42	6	PCT-US95-16626-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-08-621-081A-15
; Sequence 15, Application US/08621081A

; Patent No. 5795974

; GENERAL INFORMATION:

; APPLICANT: Cole, Barry C.

; APPLICANT: Knudson, Kevin L.

; APPLICANT: Sawitzke, Allen D.

; TITLE OF INVENTION: Mycoplasma Arthritis Superantigen

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Thorpe, No. 5795974th & Western

; STREET: 9035 South 700 East, Suite 200

; CITY: Sandy

; STATE: Utah

; COUNTRY: USA

; ZIP: 84070

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage

; COMPUTER: IBM ThinkPad 340

; OPERATING SYSTEM: MS-DOS 6.22

; SOFTWARE: Word Perfect 6.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/621,081A

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/165,038

; FILING DATE: December 10, 1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Alan J. Howarth

; REGISTRATION NUMBER: 36,553

; REFERENCE/DOCKET NUMBER: T676.CJP/U-1195

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (801)566-6633

; TELEFAX: (801)566-0750

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 48

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-621-081A-15

Query Match 64.0%; Score 28.8; DB 2; Length 48;
Best Local Similarity 82.5%; Pred. No. 0.02;
Matches 33; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 aagcttagtgatggtgatgtaaaagccaatgagc 43
|||||
Db 3 AAGCTTAGTGATGGTGTGATGATCTTCAAAAACAC 42

RESULT 2

US-08-197-770-6
; Sequence 6, Application US/08197770
; Patent No. 5514548
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Method for In Vivo Selection of
; TITLE OF INVENTION: Ligand-Binding Proteins
; NUMBER OF SEQUENCES: 23
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,770
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93 10 2484.8
; FILING DATE: 17-FEB-1993
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-197-770-6

Query Match 56.9%; Score 25.6; DB 1; Length 45;
Best Local Similarity 87.5%; Pred. No. 0.28;
Matches 28; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 aagcttagtgatggtgatgtaaaagcc 35
|||||
Db 3 AAGCTTAGTGATGGTGTGATGATGATTTCC 34

RESULT 3

US-08-197-770-8
; Sequence 8, Application US/08197770
; Patent No. 5514548
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Method for In Vivo Selection of
; TITLE OF INVENTION: Ligand-Binding Proteins
; NUMBER OF SEQUENCES: 23
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,770
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93 10 2484.8
; FILING DATE: 17-FEB-1993
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-197-770-8

Query Match 56.9%; Score 25.6; DB 1; Length 45;
Best Local Similarity 87.5%; Pred. No. 0.28;
Matches 28; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 aagcttagtgatggtgatgtaaaagcc 35
|||||
Db 3 AAGCTTAGTGATGGTGTGATGATGATTTCC 34

RESULT 4

US-08-642-255-113
; Sequence 113, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertlam I.
; REGISTRATION NUMBER: 20, 015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic"
US-08-642-255-113

Query Match 56.4%; Score 25.4; DB 2; Length 32;
Best Local Similarity 96.3%; Pred. No. 0.31;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 aagcttagtgatggtgatgtaaa 31
|||||
Db 1 AAGCTTAGTGATGGTGTGATGATGAGA 27

RESULT 5

US-08-397-633A-25
; Sequence 25, Application US/08397633A

```

Patent No. 5773577
GENERAL INFORMATION:
APPLICANT: Cappello, Joseph
TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE
TITLE OF INVENTION: OF ENZYMA TIC CROSS-LINKING
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/397,633A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-397-633A-25

Query Match          56.4%; Score 25.4; DB 2; Length 32;
Best Local Similarity 96.3%; Pred. No. 0.31;
Matches    26; Conservative   0; Mismatches      1; Indels     0; Gaps     0.

QY       5  acctagatgatgtatgtatgaa 31
         |||||||
DB        1  AGCTTAGTGATGTCGTGTCATGAGA 27

RESULT      6
US-08-197-770-2
Sequence 2, Application US/08197770
Patent No. 5514548
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A Method for In Vivo Selection of
TITLE OF INVENTION: Ligand-Binding Proteins
NUMBER OF SEQUENCES: 23
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/197,770
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 93 10 2484.8
FILING DATE: 17-FEB-1993
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 base pairs

```

```

? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
US-08-197-770-2

Query Match          55.6%; Score 25; DB 1; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      4 aagcttagtgcgtgacgtcgtatc 28
        |||
Db       3 AAGCTTACTGATGGTGATGTGATG 27

RESULT 7
US-08-197-770-12
: Sequence 12, Application US/08197770
: Patent No. 5514548
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: A Method for In Vivo Selection of
: NUMBER OF SEQUENCES: 23
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/197,770
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: EP 93 10 2484.8
: FILING DATE: 17-FEB-1993
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 54 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
US-08-197-770-12

Query Match          55.6%; Score 25; DB 1; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      4 aagcttagtgcgtgacgtcgtatc 28
        |||
Db       3 AAGCTTACTGATGGTGATGTGATG 27

RESULT 8
US-08-697-631A-12
: Sequence 12, Application US/08697631A
: Patent No. 5879673
: GENERAL INFORMATION:
: APPLICANT: Thomas, Griffith R
: TITLE OF INVENTION: NO. 5879673e1 Administration of Thrombopoietin
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 1 DNA Way
: CITY: South San Francisco
: STATE: California

```

```

: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Winpatin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/697,631A
: FILING DATE: 28-Aug-1996
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/641443
: FILING DATE: 29-Apr-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/591925
: FILING DATE: 25-Jan-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Schwartz, Timothy R.
: REGISTRATION NUMBER: 32171
: REFERENCE/DOCKET NUMBER: P0989P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650/225-7467
: TELEFAX: 650/952-9881
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 64 base pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: MOLECULE TYPE: DNA (genomic)
: US-08-697-631A-12

Query Match          54.7%; Score 24.6; DB 3; Length 64;
Best Local Similarity 87.1%; Pred. No. 0.68;
Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 tagtgatggtgatggtgatgaaagccaat 39
   | ||||| ||||| ||||| ||||| |||||
Db 17 TGTGTATGTTGATGATGATGAATGCGATAT 47

RESULT 9
US-08-697-631A-13/C
: Sequence 13, Application US/08697631A
: Patent No. 5879673
: GENERAL INFORMATION:
: APPLICANT: Thomas, Griffith R
: TITLE OF INVENTION: No. 5879673el Administration of Thrombopoietin
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 1 DNA Way
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Winpatin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/697,631A
: FILING DATE: 28-Aug-1996
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/641443
: FILING DATE: 29-Apr-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/591925
: FILING DATE: 25-Jan-1996
```

```

: ATTORNEY/AGENT INFORMATION:
: NAME: Schwartz, Timothy R.
: REGISTRATION NUMBER: 32171
: REFERENCE/DOCKET NUMBER: P0989P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650/225-7467
: TELEFAX: 650/952-9881
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 65 base pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: MOLECULE TYPE: DNA (genomic)
: US-08-697-631A-13

Query Match          54.7%; Score 24.6; DB 3; Length 65;
Best Local Similarity 87.1%; Pred. No. 0.68;
Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 tagtgatggtgatggtgatgaaagccaat 39
   | ||||| ||||| ||||| ||||| |||||
Db 49 TGTGTATGTTGATGATGATGAATGCGATAT 19

RESULT 10
US-08-697-631A-11/C
: Sequence 11, Application US/08697631A
: Patent No. 5879673
: GENERAL INFORMATION:
: APPLICANT: Thomas, Griffith R
: TITLE OF INVENTION: No. 5879673el Administration of Thrombopoietin
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 1 DNA Way
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Winpatin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/697,631A
: FILING DATE: 28-Aug-1996
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/641443
: FILING DATE: 29-Apr-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/591925
: FILING DATE: 25-Jan-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Schwartz, Timothy R.
: REGISTRATION NUMBER: 32171
: REFERENCE/DOCKET NUMBER: P0989P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650/225-7467
: TELEFAX: 650/952-9881
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 66 base pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: MOLECULE TYPE: DNA (genomic)
: US-08-697-631A-11
```


Sequence 21, Application US/08642246
Patent No. 6033654
GENERAL INFORMATION:
APPLICANT: STEDRONSKY, Erwin R.
APPLICANT: CARPELLO, Joseph
TITLE OF INVENTION: Tissue Adhesive Using Synthetic

;; TITLE OF INVENTION: Crosslinking
;; NUMBER OF SEQUENCES: 35
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;; STREET: Four Embarcadero Center, Suite 200
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: US
;; ZIP: 94111
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/642,246
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ROWLAND, Bertram I
;; REGISTRATION NUMBER: 20015
;; REFERENCE/DOCKET NUMBER: A61127-1/BIR
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-781-1989
;; TELEFAX: 415-398-3249
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 32 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; DESCRIPTION: /desc = "synthetic"
US-08-642-246-21

Query Match 54.2%; Score 24.4; DB 5; Length 32;
Best Local Similarity 96.2%; Pred. No. 0.71;
Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 gcttagtgatggtgatggtgaaga 31
|||||
Db 2 gcttagtgatggtgatggtgaaga 27

RESULT 14
PCT-US96-06229-21
;; Sequence 21, Application PC/TUS9606229
;; GENERAL INFORMATION:
;; APPLICANT: STEDRONSKY, Erwin R.
;; APPLICANT: CAPPELLO, Joseph
;; TITLE OF INVENTION: Tissue Adhesive Using Synthetic
;; TITLE OF INVENTION: Crosslinking
;; NUMBER OF SEQUENCES: 35
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;; STREET: Four Embarcadero Center, Suite 200
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: US
;; ZIP: 94111
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US96/06229
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ROWLAND, Bertram I
;; REGISTRATION NUMBER: 20015

;; REFERENCE/DOCKET NUMBER: A61127-1/BIR
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-781-1989
;; TELEFAX: 415-398-3249
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 32 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; DESCRIPTION: /desc = "synthetic"
PCT-US96-06229-21

Query Match 54.2%; Score 24.4; DB 6; Length 32;
Best Local Similarity 96.2%; Pred. No. 0.71;
Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 gcttagtgatggtgatggtgaaga 31
|||||
Db 2 gcttagtgatggtgatggtgaaga 27

RESULT 15
US-08-707-237A-83
;; Sequence 83, Application US/08707237A
;; Patent No. 5830713
;; GENERAL INFORMATION:
;; APPLICANT: Ferrari, Franco A.
;; APPLICANT: CAPPELLO, Joseph
;; APPLICANT: CRISSEMAN, John W.
;; APPLICANT: DORMAN, Mary A.
;; TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC
;; TITLE OF INVENTION: REPETITIVE DNA
;; NUMBER OF SEQUENCES: 108
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
;; STREET: Four Embarcadero Center, Suite 3400
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: United States
;; ZIP: 94111-4187
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/707,237A
;; FILING DATE: 03-SEP-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/175,155
;; FILING DATE: 29-DEC-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/053,049
;; FILING DATE: 22-APR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/609,716
;; FILING DATE: 06-NOV-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/269,429
;; FILING DATE: 09-NOV-1988
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/114,618
;; FILING DATE: 29-OCT-1987
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 06/927,258
;; FILING DATE: 04-NOV-1986
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Trecaulin, Richard F.
;; REGISTRATION NUMBER: 31,801

REFERENCE/DOCKET NUMBER: A-55186-10/WHD
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 781-1989
 TELEFAX: (415) 398-3249
 TELEX: 910 277299
 INFORMATION FOR SEQ ID NO: 83:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 31 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-08-707-237A-83

Query Match 49.8%; Score 22.4; DB 3; Length 31;
 Best Local Similarity 95.8%; Pred. No. 3.7;
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 8 ttagtgatggtgatgtaaa 31
 ||||||||||||||||
 Db 3 ttagtgatggtgatgtaaa 26

Search completed: June 4, 2000, 16:09:46
 Job time: 28065 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:41 ; Search time 1236.38 Seconds

(without alignments)
-31.472 Million cell updates/sec

Title: US-09-164-714-17

Perfect score: 40
Sequence: 1 gacggccggcgtgtatcaattgcatagcgtaagtt 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Genembl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pl4:*
11: gb_pl5:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_hum3:*
21: em_in:*
22: em_om:*
23: em_or:*
24: em_ov:*
25: em_pat:*
26: em_ph:*
27: em_pl:*
28: em_ro:*
29: em_sts:*
30: em_sy:*
31: em_un:*
32: em_vl:*
33: gb_ba1:*
34: gb_ba2:*
35: gb_ba3:*
36: gb_ba4:*
37: gb_ba5:*
38: gb_ba6:*
39: gb_ba7:*
40: gb_ba8:*
41: gb_ba9:*
42: gb_ba10:*
43: gb_ba11:*
44: gb_ba12:*

45: gb_ba13:*
46: gb_ba14:*
47: gb_ba15:*
48: gb_ba16:*
49: gb_ba17:*
50: gb_ba18:*
51: gb_ba19:*
52: gb_ba20:*
53: gb_ba21:*
54: gb_ba22:*
55: gb_ba23:*
56: gb_ba24:*
57: gb_ba25:*
58: gb_ba26:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	16.8	42.0	81	13 HUMUT815A	L31216 Human STS U
2	16.6	41.5	86	9 D78279S03	D78281 Homo sapien
3	16.2	40.5	69	12 AF158139	AF158139 Mus muscu
4	16	40.0	20	5 A93559	A93559 Sequence 4
5	16	40.0	45	14 D64066	D64066 Synthetic c
6	16	40.0	62	10 S85457	S85457 Immunoglobu
7	15.6	39.0	60	10 S69473	S69473 Igh-Ig heav
8	15.4	38.5	60	10 S75386	S75386 Smu-S gamma
9	15.2	38.0	60	40 AF064896	AF064896 Homo sapl
10	15.2	38.0	91	12 RATTCBXL	M18850 Rat T-cell
11	15	37.5	50	12 MMU41929	U41929 Mus musculu
12	15	37.5	94	5 AR017623	AR017623 Sequence
13	14.8	37.0	62	5 A84780	A84780 Sequence 25
14	14.8	37.0	66	12 S73712	S73712 TCR V1402.2
15	14.8	37.0	99	12 AF041885	AF041885 Mus muscu
16	14.6	36.5	35	5 A39021	A39021 Sequence 2
17	14.6	36.5	35	5 AR016008	AR016008 Sequence
18	14.6	36.5	39	24 E09619	E09619 PCR primer
19	14.6	36.5	60	12 M051G44	M12445 Mouse Ig ge
20	14.6	36.5	74	9 HS33VB3	Z43302 H. sapiens T
21	14.6	36.5	88	34 TCUC39755	U39755 Trypanosoma
22	14.6	36.5	89	12 MMV5M204	Z12386 M.musculus
23	14.6	36.5	100	9 HS185B3F	Z60005 H. sapiens C
24	14.4	36.0	46	5 A93653	A93653 Sequence 5
25	14.4	36.0	57	13 HUMUT5347A	L30843 Human STS U
26	14.4	36.0	65	5 A87799	A87799 Sequence 18
27	14.4	36.0	65	5 A87800	A87800 Sequence 19
28	14.4	36.0	66	12 AF158167	AF158167 Mus muscu
29	14.4	36.0	78	10 HSU5079	U5079 Human 18cla
30	14.4	36.0	93	5 AR024220	AR024220 Sequence
31	14.4	36.0	94	13 HUMUT5143A	L31061 Human STS U
32	14.2	35.5	38	5 E05849	E05849 DNA sequenc
33	14.2	35.5	49	5 AR018160	AR018160 Sequence
34	14.2	35.5	45	5 I71071	I71071 Sequence 22
35	14.2	35.5	50	12 MMU41915	U41915 Mus musculu
36	14.2	35.5	60	5 I00798	I00798 Sequence 2
37	14.2	35.5	60	5 I00800	I00800 Sequence 4
38	14.2	35.5	73	5 AR012449	AR012449 Sequence
39	14.2	35.5	73	5 AR020277	AR020277 Sequence
40	14.2	35.5	73	5 I82623	I82623 Sequence 64
41	14.2	35.5	95	10 HSSRP2S05	U72200 Human pheno
42	14	35.0	22	5 AR049652	AR049652 Sequence
43	14	35.0	31	5 I28696	I28696 Sequence 18
44	14	35.0	37	5 A42247	A42247 Sequence 35
45	14	35.0	60	10 S69475	S69475 Igh-Ig heav

ALIGNMENTS

Query Match 40.5%; Score 16.2; DB 12; Length 69;
Best Local Similarity 64.9%; Pred. No. 1.1e+04;
Matches 24; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 4 ggcgcggctgtatcaatgcatgagcggtacgtt 40
||||| | | | | | | | | | | | | | | | | |
DB 23 GGCCCGGCGAGCTCTACTTGTGTGAAGGCTCAAGCT 59

RESULT 4
A93559 20 bp DNA PAT 22-JAN-2000
LOCUS Sequence 4 from Patent WO9738105.
ACCESSION A93559
VERSION A93559.1 GI:6741764
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Baier, M. and Metzner, K.
TITLE GENOMIC NUCLEIC ACIDS, CDNA AND mRNA WHICH CODE FOR POLYPEPTIDES WITH IL-16 ACTIVITY, PROCESSES FOR THE PRODUCTION THEREOF AND THEIR USE
JOURNAL Patent: WO 9738105-A 16-OCT-1997;
FEATURES BUNDESREP DEUTSCHLAND (DE); BAIER MICHAEL (DE)
source location/Qualifiers
1..20
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 1 a 6 c 10 g 3 t
ORIGIN

Query Match 40.0%; Score 16; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacgcgcggctgtgt 16
||||| | | | | | | | | | | | | | | | | |
DB 5 GACGCGCCGGCTGTGT 20

RESULT 5
D64066 45 bp DNA SYN 21-MAY-1999
LOCUS Synthetic construct DNA for TSST-1 binding peptide, partial cds.
DEFINITION D64066
ACCESSION D64066
VERSION D64066.1 GI:1665732
KEYWORDS TSST-1 binding peptide.
SOURCE synthetic construct DNA, clone:Phage 8.
ORGANISM artificial sequence.
REFERENCE 1 (sites)
AUTHORS Sato, A., Ida, N., Fukuyama, M., Miwa, K., Kazami, J. and Nakamura, H.
TITLE Identification from a phage display library of peptides that bind to toxic shock syndrome toxin-1 and that inhibit its binding to major histocompatibility complex (MHC) class II molecules
JOURNAL Biochemistry 35 (32), 10441-10447 (1996)
MEDLINE 96339315
REFERENCE 2 (bases 1 to 45)
AUTHORS Sato, A., Ida, N., Fukuyama, M., Miwa, K., Kazami, J. and Nakamura, H.
TITLE Identification of Toxic Shock Syndrome Toxin-1 binding peptides that inhibit binding to MHC class II molecules from a phage display library
JOURNAL Unpublished (1995)
REFERENCE 3 (bases 1 to 45)
AUTHORS Sato, A.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-1995) to the DBJ/EMBL/Genbank databases. Atsushi Sato, Biomolecular Engineering Research Institute, Department of Molecular Biology, Furuhashi 6-2-3, Suita, Osaka 565-0874, Japan (Tel:06-872-8204, Fax:06-872-8219)

FEATURES Location/Qualifiers
source 1..45
/organism="synthetic construct"
/db_xref="taxon:32630"
/clone="Phage 8"
/note="synonym:artificial sequence"
CDS <1..>45
/function="TSST-1 binding"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="TSST-1 binding peptide"
/protein_id="BAI10941.1"
/db_xref="GI:1665733"
/translation="WDAMTWNRSEFH"
BASE COUNT 8 a 4 c 17 g 16 t
ORIGIN

Query Match 40.0%; Score 16; DB 14; Length 45;
Best Local Similarity 68.8%; Pred. No. 1.3e+04;
Matches 22; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 9 ggcgtgtatcaatgcatgagcggtacgtt 40
||| | | | | | | | | | | | | | | | |
DB 2 GGGATGCGATGTAFTGGATTGGCGGAGTGTT 33

RESULT 6
S85457/c 62 bp DNA PRI 07-MAY-1993
LOCUS Immunoglobulin heavy chain complementarity determining region (CDR)-III {3' variable, diversity, and 5' joining regions} [human, B cell precursor acute lymphoblastic leukemia cells, Genomic, 62 nt].
DEFINITION S85457
ACCESSION S85457
VERSION S85457.1 GI:246644
KEYWORDS
SOURCE human B cell precursor acute lymphoblastic leukemia cells.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 62)
AUTHORS Kiyoi, H., Naoe, T., Horibe, K. and Ohno, R.
TITLE Characterization of the immunoglobulin heavy chain complementarity determining region (CDR)-III sequences from human B cell precursor acute lymphoblastic leukemia cells
JOURNAL J. Clin. Invest. 89 (3), 739-746 (1992)
MEDLINE 92176356
REMARK Genbank staff at the National Library of Medicine created this entry [NCBI g1bbsq 85457] from the original journal article.
This sequence comes from fig 1.
FEATURES Location/Qualifiers
1..62
source /organism="Homo sapiens"
/db_xref="taxon:9606"
1..62
gene /partial
/gene="Immunoglobulin heavy chain complementarity determining region (CDR)-III"

BASE COUNT 10 a 17 c 23 g 12 t
ORIGIN

Query Match 40.0%; Score 16; DB 10; Length 62;
Best Local Similarity 68.8%; Pred. No. 1.3e+04;
Matches 22; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 6 cccgcgcgtgtatcaatgcatgagcggtac 37
||| | | | | | | | | | | | | | | | |
DB 47 CCAGGGGTGCAACCACTTGTCTAAGAGGTAA 16

RESULT 7
LOCUS S69473 60 bp DNA PRI 23-SEP-1994
DEFINITION IGH-Ig heavy chain [clone PSC1.1-11, mu to epsilon switch circle, recombination breakpoint] [human, B lymphocytes, Extrachromosomal, 60 nt].
ACCESSION S69473 GI:546126
VERSION S69473.1 GI:546126
KEYWORDS
SOURCE human B lymphocytes.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 60)
AUTHORS Zhang, K., Mills, F.C. and Saxon, A.
TITLE Switch circles from IL-4-directed epsilon class switching from human B lymphocytes. Evidence for direct, sequential, and multiple step sequential switch from mu to epsilon Ig heavy chain gene
JOURNAL J. Immunol. 152 (7), 3427-3435 (1994)
MEDLINE 94194102
REMARK Genbank staff at the National Library of Medicine created this entry [NCBI g1bbsq 146011] from the original journal article. This sequence comes from Fig. 3A.
FEATURES
source Location/Qualifiers
1..60
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..60
/partial
/gene="IGH"
/note="Ig heavy chain"
BASE COUNT 13 a 12 c 23 g 12 t
ORIGIN
Query Match 39.0%; Score 15.6; DB 10; Length 60;
Best Local Similarity 81.8%; Pred. No. 2e+04;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 gggccggcgtgtatcaatg 25
||| ||||| ||| |||
Db 33 GGCTGGCTGGAGTGAACCTGG 54
RESULT 8
LOCUS S75386 60 bp DNA PRI 27-JUL-1995
DEFINITION Smu-S gamma 1 switch region: Ig VDJ (alternatively spliced, switch recombination sites) [human, isotype-switched B-lymphocytes, Genomic, 60 nt].
ACCESSION S75386
VERSION S75386.1 GI:914260
KEYWORDS
SOURCE human isotype-switched B-lymphocytes.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 60)
AUTHORS Zhang, K., Chean, H.K. and Saxon, A.
TITLE Secondary deletional recombination of rearranged switch region in Ig isotype-switched B cells. A mechanism for isotype stabilization
JOURNAL J. Immunol. 154 (5), 2237-2247 (1995)
MEDLINE 95173430
REMARK Genbank staff at the National Library of Medicine created this entry [NCBI g1bbsq 160356] from the original journal article. This sequence comes from Fig. 4B.
COMMENT Region: Smu-S gamma 1 switch region.
FEATURES
source Location/Qualifiers
1..60
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..60
/partial
gene

BASE COUNT 10 a 13 c 24 g 13 t
ORIGIN
Query Match 38.5%; Score 15.4; DB 10; Length 60;
Best Local Similarity 76.0%; Pred. No. 2.4e+04;
Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 gggccggcgtgtatcaatgcat 28
||| ||||| ||| |||
Db 33 GGCTGGCTGGAGTGAACCTGG 57
RESULT 9
LOCUS AF064896 60 bp mRNA PRI 03-MAR-1999
DEFINITION Homo sapiens clone 11-4 immunoglobulin heavy chain alpha VDJ region mRNA, partial cds.
ACCESSION AF064896
VERSION AF064896.1 GI:4321882
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 60)
AUTHORS Rowley, A.H., Eckerley, C.A., Shulman, S.T. and Baker, S.C.
TITLE Evidence of an antigen-driven IgA immune response in the vascular wall in acute Kawasaki Syndrome
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 60)
TITLE Rowley, A.H., Eckerley, C.A., Shulman, S.T. and Baker, S.C.
JOURNAL Direct Submission
TITLE Submitted (14-MAY-1998) Pediatrics, Loyola University, 2160 S First Avenue, Maywood, IL 60153, USA
FEATURES
source Location/Qualifiers
1..60
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="11-4"
/tissue-type="vascular"
/note="from aorto-iliac junction of a 10-year-old child with fatal acute Kawasaki Syndrome"
1..>60
/note="includes IGHV6 gene"
/codon_start=1
/product="immunoglobulin heavy chain alpha VDJ region"
/protein_id="AAD15860.1"
/db_xref="GI:4321883"
/translation="PDDTAVYYCARGVGEFDYWG"
BASE COUNT 11 a 13 c 23 g 13 t
ORIGIN
Query Match 38.0%; Score 15.2; DB 40; Length 60;
Best Local Similarity 63.9%; Pred. No. 2.9e+04;
Matches 23; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 gacggccggcgtgtatcaatgcatgagcgagta 36
||| ||| ||| ||| ||| |||
Db 4 GACGACACGCGTGTGTACTGTGCAAGGGGGTA 39
RESULT 10
LOCUS RATTCBXL 91 bp mRNA ROD 27-APR-1993
DEFINITION Rat T-cell receptor active beta-chain V-region (V-J-beta2.2) mRNA, partial cds, clone TRB89.
ACCESSION M18850
VERSION M18850.1 GI:207243
KEYWORDS J-region; T-cell receptor; T-cell receptor beta-chain; V-region; processed gene.

SOURCE Rat (6- to 8-week-old) thymocyte, cDNA to mRNA, clone TR889.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE 1 (bases 1 to 91)
 AUTHORS Morris, M., Barclay, A.N. and Williams, A.F.
 TITLE Analysis of T cell receptor Beta chains in rat thymus, and rat C-alpha and C-beta sequences
 JOURNAL Immunogenetics 27, 174-179 (1988)
 MEDLINE 88113841
 FEATURES
 source Location/Qualifiers
 1. . 91
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 BASE COUNT 19 a 19 c 33 g 20 t
 ORIGIN
 Query Match 38.0%; Score 15.2; DB 12; Length 91;
 Best Local Similarity 63.9%; Pred. No. 2.9e+04;
 Matches 23; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 5 gcccgagctgtatcataatggcagcgaagt 40
 1 ||||| ||| ||| ||| |||
 Db 45 GACCGGCGACGCTACTTGTGTCAGGCTCAAGCT 80
 RESULT 11
 LOCUS MMU41929 50 bp DNA ROD 05-JAN-1996
 DEFINITION Mus musculus recombination between immunoglobulin heavy chain and c-myc.
 VERSION 041929
 KEYWORDS U41929.1 GI:1147663
 SOURCE house mouse strain-BALB/cAn.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 50)
 AUTHORS Muller, J.R.
 TITLE Direct Submission
 JOURNAL Submitted (05-DEC-1995) Jorgen R. Muller, Lab of Genetics, NIH/NCI, Bldg. 37, Room 2B09, 37 Convent Dr., Bethesda, MD 20892-4255, USA
 FEATURES
 source Location/Qualifiers
 1. . 50
 /organism="Mus musculus"
 /strain="BALB/cAn"
 /db_xref="taxon:10090"
 /chromosome="T(12;15)"
 /map="T(12;15d2)"
 /tissue_type="Spleen 7 days post pristane"
 /dev_stage="7 days post pristane"
 BASE COUNT 10 a 21 c 13 g 6 t
 ORIGIN
 Query Match 37.5%; Score 15; DB 12; Length 50;
 Best Local Similarity 67.7%; Pred. No. 3.5e+04;
 Matches 21; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 4 gcccgagctgtatcataatggcagcga 34
 1 ||| ||||| || ||||| |||
 Db 31 GTCCAGGCTGTAATGTTGGCTGGGCTG 1
 RESULT 12
 LOCUS AR017623 94 bp DNA PAT 05-DEC-1998
 DEFINITION Sequence 24 from patent US 5780228.
 ACCESSION AR017623
 VERSION AR017623.1 GI:3973226
 KEYWORDS

SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 94)
 AUTHORS Parma, D.H., Hicke, B., Bridonau, P. and Gold, L.
 TITLE High affinity nucleic acid ligands to lectins
 JOURNAL Patent: US 5780228-A 24 14-JUL-1998;
 FEATURES
 source Location/Qualifiers
 1. . 94
 /organism="unknown"
 BASE COUNT 32 a 21 c 34 g 7 t
 ORIGIN
 Query Match 37.5%; Score 15; DB 5; Length 94;
 Best Local Similarity 61.5%; Pred. No. 3.5e+04;
 Matches 24; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 2 accggccggctgtatcataatggcagcgaagt 40
 1 ||||| ||||| ||| ||||| ||| |||
 Db 35 ATGGCCGGGCGGAGAGTGGGTGGCAGGAGCCAGCT 73
 RESULT 13
 LOCUS A84780 62 bp DNA PAT 21-JAN-2000
 DEFINITION Sequence 25 from patent WO9844135.
 ACCESSION A84780
 VERSION A84780.1 GI:6733648
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 62)
 AUTHORS Hegemann, J. and Zimmermann, F.K.
 TITLE METHOD FOR SCREENING ANTIMYCOTICALLY ACTIVE SUBSTANCES
 JOURNAL Patent: WO 9844135-A 08-OCT-1998;
 HEGEMANN JOHANNES (DE); ZIMMERMANN FRIEDRICH KARL (DE)
 FEATURES
 source Location/Qualifiers
 1. . 62
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 17 a 15 c 12 g 18 t
 ORIGIN
 Query Match 37.0%; Score 14.8; DB 5; Length 62;
 Best Local Similarity 73.1%; Pred. No. 4.3e+04;
 Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 14 ggtacattggcagcgaagt 39
 1 ||||| ||| ||||| ||| |||
 Db 30 GTTATCTATTAGCATAGCCACTAGT 55
 RESULT 14
 LOCUS S73712 66 bp mRNA ROD 02-MAR-1995
 DEFINITION TCR V14J2.2 beta-myelin basic protein-specific T-cell receptor beta chain VDJ junction region (rats, Lewis, T cells, mRNA Partial), 66 nt].
 ACCESSION S73712
 VERSION S73712.1 GI:688024
 KEYWORDS
 SOURCE Rattus sp. Lewis T cells.
 ORGANISM Rattus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE 1 (bases 1 to 66)
 AUTHORS Sun, D., Shah, R. and Coleclough, C.
 TITLE Repertoire of rat MBP-reactive T cells: DNA sequencing analysis further demonstrates the clonal heterogeneity of rat T cells

reactive against encephalitogenic epitopes
Cell. Immunol. 156 (2), 389-401 (1994)
94298095
MEDLINE
Genbank staff at the National Library of Medicine created this
entry [NCBI 91bbq 156543] from the original journal article.
REMARK This sequence comes from Fig. 5.
FEATURES
SOURCE
1..66
/organism="Rattus sp."
/db_xref="taxon:10118"
1..66
/partial
/gene="TCR V14J2.2&br;"
/note="myelin basic protein-specific T-cell receptor beta
chain VDJ junction region"
1..66
/partial
/gene="TCR V14J2.2&br;"
/note="This sequence comes from Fig. 5; Protein sequence
is in conflict with the conceptual translation;
mismatches(15[E->Q],17[S->T],19[L->M])"
/codon_start=1
/product="myelin basic protein-specific T-cell receptor
beta chain VDJ junction region"
/protein_id="AAB31305.1"
/db_xref="GI:688025"
/translation="CASHRRATGQLYFEGSKLTVL"
BASE COUNT 16 a 20 c 19 g 11 t
ORIGIN
Query Match 37.0%; Score 14.8; DB 12; Length 66;
Best Local Similarity 64.7%; Pred. No. 4.3e+04;
Matches 22; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 7 ccggcgtgatacattgacataggcgytaagt 40
||||| | | | | | | | | | | | | | | | | |
DB 23 CCGGGCAGCTTACTTGGCCAGGCACCAAGAT 56
RESULT 15
AF041885 99 bp mRNA ROD 20-JAN-1998
LOCUS Mus musculus uninfected mouse C seg 1, T cell receptor beta chain
DEFINITION mRNA, partial cds.
ACCESSION AF041885
VERSION AF041885
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 99)
AUTHORS Gomez,F.J., Cain,J.A., Gibbons,R., Allendeofer,R. and Deepe,G.S.
TITLE V beta 4 T cells exert protection in pulmonary experimental
histoplasmosis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 99)
AUTHORS Gomez,F.J., Cain,J.A., Gibbons,R., Allendeofer,R. and Deepe,G.S.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-1998) Internal Medicine, Division of Infectious
Diseases, University of Cincinnati, 231 Bethesda Ave. ML 560,
Cincinnati, OH 45267-0560, USA
FEATURES
SOURCE
1..99
/organism="Mus musculus"
/strain="C57BL/6"
/isolate="uninfected control, mouse C seg 1"
/db_xref="taxon:10090"
/tissue_type="lung"
/note="type: jbeta2.2"
<1..15
/standard_name="V beta 4"

CDS
<1..>99
/note="CDR3 region"
/codon_start=1
/product="T cell receptor beta chain"
/protein_id="AAB97192.1"
/db_xref="GI:2795993"
/translation="CASSQGLGNTGQLYFEGSKLTVLEDLRNVT"
16..27
/standard_name="N sequences and D beta 2"
28..76
/standard_name="J beta 2.2"
77..>99
/standard_name="C beta"
BASE COUNT 24 a 23 c 33 g 19 t
ORIGIN
Query Match 37.0%; Score 14.8; DB 12; Length 99;
Best Local Similarity 64.7%; Pred. No. 4.3e+04;
Matches 22; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 7 ccggcgtgatacattgacataggcgytaagt 40
||||| | | | | | | | | | | | | | | | | |
DB 32 CCGGGCAGCTTACTTGGTGAAGGCTTAAGCT 65
Search completed: June 4, 2000, 16:05:43
Job time: 27895 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:18 ; Search time 322.35 Seconds

(without alignments)
31.046 Million cell updates/sec

Title: US-09-164-714-17

Perfect score: 40
Sequence: 1 gagcgccggcgtgtatcaattgcatagcggttaagtt 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16.8	42.0	99	1 V39623	PRF1 DNA oligonuc
2	16	40.0	28	1 Q93906	Human mismatch rep
3	16	40.0	48	1 V10672	Mouse Fas ligand-1
4	16	40.0	60	1 V40188	Escherichia coli B
5	16	40.0	83	1 V99110	DNA methyltransfer
6	15.6	39.0	51	1 X15369	Mutagenic HSV-1 th
7	15.4	38.5	30	1 T59426	Human papillomavir
8	15.4	38.5	54	1 Q3546	Downstream sequenc
9	15	37.5	58	1 V28977	Human MGMT with ma
10	15	37.5	94	1 T57700	Wheat germ aggluti
11	14.8	37.0	62	1 V64262	PCR primer YPL242c
12	14.6	36.5	39	1 T02866	Fungal 18S rRNA en
13	14.6	36.5	87	1 T36186	Tagged protein DET
14	14.6	36.5	87	1 T37293	Chicken Src primer
15	14.6	36.5	87	1 T51211	Primer #1 for chic
16	14.6	36.5	87	1 T68704	PCR primer for chl
17	14.4	36.0	61	1 T23877	Human gene signatu
18	14.4	36.0	63	1 V76296	Staphylococcus aur
19	14.4	36.0	97	1 T58032	P-selectin SELEX 2
20	14.2	35.5	38	1 O64620	Protein H gene fra
21	14.2	35.5	47	1 V07587	Nucleotide sequenc
22	14.2	35.5	48	1 Q34008	Sequence upstream
23	14.2	35.5	60	1 T18142	Ribonucleotide tal
24	14.2	35.5	73	1 T71390	Red blood cell gho
25	14	35.0	22	1 T01364	Human growth hormo
26	14	35.0	22	1 V12409	Growth hormone rec
27	14	35.0	22	1 V33803	Human growth hormo
28	14	35.0	31	1 T46172	Primer for fusion
29	14	35.0	37	1 T34354	Chymosin specific
30	14	35.0	55	1 Q33999	Sequence upstream
31	14	35.0	61	1 X33461	Thermal hysteresis
32	14	35.0	63	1 T04873	Silkworm prophenol
33	14	35.0	96	1 T58022	P-selectin SELEX 2
34	13.8	34.5	18	1 Q12438	Probe HPV15 for de

35	13.8	34.5	25	1 Q12443	Primer HPV15.1 for
36	13.8	34.5	36	1 X04271	ITV linker regions
37	13.8	34.5	55	1 T59173	DNA polymerase lig
38	13.8	34.5	60	1 V40207	Escherichia coli B
39	13.8	34.5	77	1 T88485	Secretory phosphol
40	13.8	34.5	94	1 Q89695	Vector back-end ol
41	13.8	34.5	94	1 Q89695	Vector back-end ol
42	13.6	34.0	36	1 V29556	Dengue virus type
43	13.6	34.0	55	1 T34005	Primer for amplifi
44	13.6	34.0	66	1 Q56624	Fragment of homolo
45	13.6	34.0	66	1 T00547	Equine herpesvirus

ALIGNMENTS

RESULT 1
V39623
ID V39623 standard; CDNA; 99 BP.
AC V39623;
DE 28-SEP-1998 (first entry)
KW PRF1 DNA oligonucleotide seq ID NO:141.
KW Mass spectrometry; diagnosis; detection; biological sample; infection;
KW genetic disease; chromosomal abnormality; identification; heredity;
KW pathogenic organism; telomerase activity; oncogene mutation;
KW cancer-specific sequence; primer; ss.
OS Synthetic.
PN WO9820166-A2.
PD 14-MAY-1998.
PF 06-NOV-1997; U20444.
PI 08-OCT-1997; US-947801.
PR 06-NOV-1996; US-744481.
PR 06-NOV-1996; US-744590.
PR 06-NOV-1996; US-746036.
PR 06-NOV-1996; US-746055.
PR 23-JAN-1997; US-786988.
PR 23-JAN-1997; US-787639.
PR 19-SEP-1997; US-933792.
PA (SEQID-) SEQUENOM INC.
PI Braun A, Damhofer-Demar B, Fu D, Higgins GS, Jurinke C,
PI Koster H, Little DP, Lough DW, Siebert CW, Tang K,
PI Van Den Boom D, Xiang G;
DR WPI; 98-286975/25.
PT Sequencing nucleic acid by mass spectrometric analysis - for
PT detecting nucleic acids, telomerase activity, oncogene mutations, or
PT cancer-specific sequences, for diagnosis of disease
PS Example 8; Page 283; 478bp; English.
CC A process has been developed for determining the sequence of a target
CC nucleic acid. The process comprises: (i) generating at least two
CC fragments (F) from the target nucleic acid; and (ii) analysing F by
CC mass spectrometry (MS). The sequences in V39483 to V39592 are
CC specifically claimed primers for use in the mass spectrometric analysis
CC of the above process. The process is used to detect genetic diseases
CC (e.g. haemophilia, thalassemia, Duchenne muscular dystrophy, Alzheimer's
CC disease, cystic fibrosis and many others) or chromosomal abnormalities
CC (or predisposition); infections and cancers; also for establishing
CC identity and heredity. Particular applications are diagnosis of
CC neuroblastoma, detecting telomerase, determining family relationships
CC and HLA compatibility, and in genetic fingerprinting. Compared with
CC known methods using MS, this process requires fewer specific reagents
CC and is better suited to automation. Extended primers are shorter;
CC primer annealing is more efficient and the process allows detection of
CC many sequences simultaneously. The present sequence represent an
CC oligonucleotide used in an example from the present invention.
SQ Sequence 99 BP; 24 A; 34 C; 20 G; 21 T;
Query Match 42.0%; Score 16.8; DB 1; Length 99;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
OY 5 gcccgagctgtatcaattgcatagcggttaagtt 40
|| ||||| ||||| ||| || ||

Db 59 GCATGGCTGCAGCCATTGGCACTGGCCGTCGTTT 94

RESULT 2

ID 093906/c

AC 093906; standard; DNA; 28 BP.

DT 15-JAN-1996 (first entry)

DE Human mismatch repair pathway gene MSH2 exon primer, 16066.

KW Mismatch repair; MSH2; primer; identification; defect; alteration;

KW cancer; tumour; vaccine; ss.

OS Homo sapiens.

PN W09514085-A2.

PD 26-MAY-1995.

PF 17-NOV-1994; U13385.

PR 17-NOV-1993; US-154792.

PR 07-DEC-1993; US-163449.

PR 13-JUN-1994; US-259310.

PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.

PA (DAND) DANA FARBEN CANCER INST.

PI Fisher R, Kolodner RD, Reenan RAG;

DR WPI: 95-200377/26.

PT Determining alteration in human mismatch repair pathways - used in

PT the diagnosis, prognosis and therapy of cancers and in screening

PT assays

PS Claim 30; Page 176; 256pp; English.

CC 093906 is an oligonucleotide primer used to amplify an exon

CC of the human mismatch repair pathway gene MSH2. Defects or

CC alterations in such a gene result in the accumulation of unstable

CC repeated DNA sequences, a feature of a number of different cancers.

CC The identification of a defect in the mismatch repair pathway can

CC be diagnostic of a predisposition to cancer and prognostic for a

CC particular mammalian cancer e.g colorectal, ovarian, endometrial

CC (uterine), renal, bladder, skin, rectal and bowel. The nucleotide

CC sequences and polypeptides of the hMSH2 gene may also be used for

CC therapy and in vaccines.

SQ Sequence 28 BP; 8 A; 8 C; 5 G; 7 T;

Query Match 40.0%; Score 16; DB 1; Length 28;

Best Local Similarity 79.2%; Pred. No. 2.1e+02;

Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 9 ggcgcgtgatacattgcatgacgc 32

Db 24 GGACGCTTTCATGACATGACATGCGC 1

RESULT 3

ID V10672 standard; cDNA; 48 BP.

AC V10672;

DT 02-JUL-1998 (first entry)

DE Mouse Fas ligand-like protein oligonucleotide #3.

KW Fas antigen; ligand-like; apoptosis-inducing activity; therapeutic;

KW prophylactic; cancer; viral infection; hepatitis; nephritis; influenza;

KW bone disease; atherosclerosis; inflammation; neurodegenerative disease;

KW ischaemic disease; autoimmune disease; diabetes; cardiomyopathy;

KW ulcerative colitis; primer; ss.

OS Synthetic.

OS Mus sp.

PN W09803648-A1.

PD 29-JAN-1998.

PF 17-JUL-1997; J02480.

PR 31-JAN-1997; JP-019330.

PR 19-JUL-1996; JP-191204.

PR 09-AUG-1996; JP-211695.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Hikichi Y, Nishi K, Shintani Y;

DR WPI: 98-120773/11.

PT Fas ligand-like protein with apoptosis inducing activity - useful as

PT therapeutic or prophylactic agent for cancer, viral infection,

PT Helicobacter pylori infection, invasive staphylococci, hepatitis,

PT nephritis, bone disease, atherosclerosis or pain

PS Example 2; Page 152; 120pp; English.

CC Oligonucleotides V10670-V10674 are used in a method to clone a novel

CC mouse Fas ligand-like protein with apoptosis inducing activity. This

CC protein can be used as a therapeutic and/or prophylactic agent for

CC cancer, viral infection, Helicobacter pylori infection, invasive

CC staphylococci, hepatitis, nephritis, bone disease, atherosclerosis or

CC pain. The protein can also be used in a method to treat acquired immune

CC deficiency syndrome (AIDS), joint tissues in rheumatism, inflammation,

CC hepatitis, neurodegenerative disease due to osteomyelodysplasia,

CC ischaemic disease, autoimmune disease, diabetes, diabetic complications,

CC cardiomyopathy, influenza, glomerulonephritis or ulcerative colitis.

SQ Sequence 48 BP; 9 A; 13 C; 17 G; 9 T;

Query Match 40.0%; Score 16; DB 1; Length 48;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacggcccgagctgt 16

Db 33 GACGCCCGGCGCTGCT 48

RESULT 4

ID V40188/c

AC V40188; standard; DNA; 60 BP.

DT 08-OCT-1998 (first entry)

DE Escherichia coli B1AP gene codon usage library construction oligo 9.

KW Escherichia coli; B1AP; codon usage library; construction; evolution;

KW protein engineering; industrial; pharmaceutical; recombination;

KW selection; promoter; ss.

OS Synthetic.

OS Escherichia coli.

PN W09827230-A1.

PD 25-JUN-1998.

PF 17-DEC-1997; U24239.

PR 18-DEC-1996; US-769062.

PA (MAXY-) MAXYGEN INC.

PI Patten PA, Stemmer WPC;

DR WPI: 98-362797/31.

PT Library of recombinant DNA molecules and screening for improved

PT properties such as activity, expression, solubility or specificity

PT Example; Page 61; 122pp; English.

PS A method has been developed for evolving a protein encoded by a DNA

CC substrate molecule. The method comprises: (a) digesting at least a first

CC and second DNA substrate molecule with restriction endonuclease, where

CC the first and second substrate molecules differ from each other in at

CC least one nucleotide; (b) ligating the mixture to generate a library of

CC recombinant DNA molecules; (c) screening or selecting the products of

CC (b) for a desired property; and (d) recovering a recombinant DNA

CC substrate molecule encoding an evolved protein. The method can be used

CC for evolving proteins to provide improved properties such as expression,

CC folding, solubility, activity, substrate specificity, shelf-life and

CC serum stability or circulation half-life. V40180 to V40235 represent

CC oligonucleotides used in the construction of a codon usage library in an

CC example from the present invention, describing a strategy to evolve B1AP

CC from Escherichia coli.

SQ Sequence 60 BP; 9 A; 12 C; 22 G; 13 T;

Query Match 40.0%; Score 16; DB 1; Length 60;

Best Local Similarity 64.3%; Pred. No. 2.4e+02;

Matches 18; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 7 ccggcgtggtcaattgcatgacgcg 34

Db 44 CCAGRCRATATCATGACATGACCGG 17

RESULT 5

```

V99110
ID V99110 standard; DNA; 83 BP.
AC V99110:
DE 09-MAR-1999 (first entry)
DE DNA methyltransferase genomic fragment.
KW DNA methyltransferase; DNA Methylase; antisense oligonucleotide; human;
KW cellular growth; tumour growth inhibition; silenced gene activation;
KW beta thalassemia; sickle cell anemia; ss.
OS Homo sapiens.
PN MO9854313-A2.
PD 03-DEC-1998.
PF 29-MAY-1998; IB1107.
PR 17-DEC-1997; US-069865.
PR 30-MAY-1997; US-866340.
PA (UYMC-) UNIV MCGILL.
PI Bigey P, Ramchandani S, Szyf M;
DR WPI: 99-059833/05.
PT New DNA methyltransferase nucleotide sequences - used particularly
PT to develop antisense oligonucleotides for diagnostic and therapeutic
PT purposes, particularly for inhibiting tumour growth
PS Claim 1, Page 52; 108pp; English.
CC V99091-128 represent DNA methyltransferase (DNA Methylase) genomic
CC fragments. The sequences together sequentially produce the full
CC length sequence V99128. The sequences can be used to derive antisense
CC oligonucleotides which inhibit DNA Methylase expression. The antisense
CC oligonucleotides can be used in investigating the role of DNA Methylase
CC in cellular growth. They can be administered at different points in
CC the cell cycle, or in conjugation with promoters or inhibitors of cell
CC growth to determine the role of DNA Methylase in the growth of the cell
CC type of interest. The antisense oligonucleotides can also be used for
CC inhibiting tumour growth in a mammal, or to activate silenced genes to
CC provide a missing gene function. This ameliorates disease symptoms,
CC e.g. in beta thalassemia and sickle cell anemia. The antisense
CC oligonucleotides can also be used as an analytical and diagnostic tools
CC and a potentiators of transgenic plant and animal studies.
SQ Sequence 83 BP; 24 A; 18 C; 21 G; 20 T;

Query Match 40.0%; Score 16; DB 1; Length 83;
Best Local Similarity 68.8%; Pred. No. 2.5e+02;
Matches 22; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 acggccggcgctgataatgcatagcg 33
DB 25 ACAGCCCCGGTGTCTACTTAGAAAGGG 56

RESULT 6
X15369
ID X15369 standard; DNA; 51 BP.
AC X15369;
DE 22-JUN-1999 (first entry)
DE Mutagenic HSV-1 thymidine kinase gene primer MB127.
KW HSV-1; thymidine kinase; mutation; DRH nucleoside binding site; enzyme;
KW pathogen; tumour; hyperkeratosis; psoriasis; prostate hypertrophy;
KW hyperthyroidism; endocrinopathy; autoimmune disease; allergy; restenosis;
KW viral disease; AIDS; hepatitis; parasite; bacterial infection; ss.
OS Synthetic.
OS Herpes simplex virus.
PN US5877010-A.
PD 02-MAR-1999.
PF 02-MAY-1995; 432871.
PR 02-MAY-1995; US-432871.
PR 02-MAY-1994; US-237592.
PA (UNIV ) UNIV WASHINGTON.
PI Black ME, Loeb LA;
DR WPI: 99-189650/16.
PT New Herpesviridae thymidine kinase mutant nucleic acids - used to
PT develop products for treating e.g. tumours, autoimmune diseases,
PT allergies, restenosis or viral, bacterial or parasitic diseases
PT Example 8; Column 40; 72pp; English.
CC This sequence represents a primer used to construct a mutated herpes
CC simplex virus type 1 (HSV-1) thymidine kinase (TK) gene with mutations

```

```

CC In codons 159-161 and 168-170. The invention relates to the generation
CC of novel HSV-1 TK gene with a mutation upstream, within or downstream
CC from a DRH nucleoside binding site. The TK enzymes can be used for
CC inhibiting pathogenic agents, e.g. tumours, hyperkeratosis, psoriasis,
CC prostate hypertrophy, hyperthyroidism, endocrinopathies, autoimmune
CC diseases, allergies, restenosis, viral diseases such as AIDS, hepatitis,
CC intracellular parasitic diseases or bacterial infection.
SQ Sequence 51 BP; 8 A; 9 C; 20 G; 5 T;

Query Match 39.0%; Score 15.6; DB 1; Length 51;
Best Local Similarity 53.8%; Pred. No. 3.5e+02;
Matches 21; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 2 acggccggcgctgataatgcatagcggaagtc 40
DB 8 ACCGCCGCCGCCGGTACGANNNNNNNGCGGATGAT 46

RESULT 7
T59426
ID T59426 standard; DNA; 30 BP.
AC T59426;
DE 26-MAY-1997 (first entry)
DE Human papillomavirus 6b early ORF E2 gene forward primer.
KW Human papillomavirus; HPV; polyprotein; E2 gene;
KW early open reading frame; vaccine; wart; condylomata acuminata;
KW cervical cancer; polymerase chain reaction; PCR; primer; ss.
OS Synthetic.
PN WO9705164-A1.
PD 13-FEB-1997.
PF 26-JUL-1996; AU0473.
PR 27-JUL-1995; AU-004439.
PA (CSLC-) CSL LTD.
PA (UYOU ) UNIV QUEENSLAND.
PI Cox JC, Edwards SJ, Frazer I, Margetis MB, McMillan NAJ;
PI Moloney MBH, Webb EA, Williams MP;
DR WPI: 97-145619/13.
PT Papilloma virus early open reading frame polyprotein constructs -
PT useful for treating, e.g. genital warts and cervical cancer
PS Example 1; Page 18; 62pp; English.
CC PCR primers (T59426-27) were designed to amplify the E2 gene
CC of human papillomavirus 6b (HPV6b), incorporating a 5' SmaI site
CC and 3' XbaI, NcoI, KpnI and SacI sites, and using the entire genome
CC of HPV6b in pBR322 as template. The HPV6b E2b gene was similarly
CC amplified (see also T59428-29). The amplified genes were cloned
CC into an HPV6b E6/E4 cassette (see also T59412-15) in pSP70. The
CC resulting cassettes (see also T59430-35) can be cloned into
CC pCEX-STOP to allow prodn. of HPV6b polyproteins (W01808-13) in E.
CC coli for use as vaccines.
SQ Sequence 30 BP; 8 A; 9 C; 11 G; 2 T;

Query Match 38.5%; Score 15.4; DB 1; Length 30;
Best Local Similarity 76.0%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 cggccggcgctgataatgca 27
DB 2 CGCCCGCGGATGAGCAATAGCCA 26

RESULT 8
Q33546
ID Q33546 standard; DNA; 54 BP.
AC Q33546;
DE 02-FEB-1993 (first entry)
DE Downstream sequence of microsatellite from clone AG1A233.
KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
KW genetic mapping; traits; amplification; ss.
OS Bos taurus.
PN WO9213102-A.
PD 06-AUG-1992.

```



```

RESULT 11
ID V64262 standard; DNA; 62 BP.
AC V64262;
DE 25-JAN-1999 (first entry)
DE PCR primer YPL242c-S2.
KM Antimycotic agent; target; medicine; infection; veterinary; fungicide;
KM immunodepression; preservative; food industry; fungi; PCR primer; ss.
OS Synthetic.
PN WO9844135-A2.
PD 08-OCT-1998.
PR 02-APR-1998: E01904.
PR 02-APR-1997: DE-013572.
PA (FARN) HOECHST AG.
PI Enlihan K, Feldmann H, Hegemann J, Hinnen A, Koetter P,
PI Kramer W, Munder T, Rose M, Schuster T, Zimmermann FK;
DR WPI: 98-557125/47.
PT Identification of antimycotic agents using essential fungal proteins
PT or genes as targets - useful, e.g. for potential clinical, human or
PT veterinary medicine, for treatment of existing infections and for
PT prevention of these in immune depressed subjects
PS Disclosure; Page 38; 76pp; German.
CC V64259-V64274 are primers used in a method for the identification of
CC antimycotic agents using as a target a nucleic acid which controls an
CC essential protein of Saccharomyces cerevisiae or from other species of
CC Mycelles. Such agents are potentially useful clinically, in human or
CC veterinary medicine, for treating existing infections and for preventing
CC them in immune-depressed subjects (those with human immune deficiency
CC virus infection or diabetes), also as fungicides and preservatives for
CC foods and body care products. The agents are used to identify equivalent
CC genes in other fungi, specifically Candida albicans or Aspergillus
CC fumigatus, and equivalent human, animal and plant genes, and also for
CC identification of antimycotic agents.
SQ Sequence 62 BP; 17 A; 15 C; 12 G; 18 T;

Query Match 37.0%; Score 14.8; DB 1; Length 62;
Best Local Similarity 73.1%; Pred. No. 8e+02;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 14 ggtacattgcatagcggttaagt 39
DB 30 GTATCTATTAGCATAGCCACTAGT 55

RESULT 12
ID T02866 standard; DNA; 39 BP.
AC T02866;
DE 14-MAR-1996 (first entry)
DE Fungal 18S rRNA encoding DNA amplification/sequencing primer, UFI.
KM Polymerase chain reaction; primer: ribosomal RNA; amplification;
KM sequencing; Matsutake mushroom; ss.
OS Agaricus bisporus.
PN J07177889-A.
PD 18-JUL-1995.
PR 22-DEC-1993: 346106.
PR 22-DEC-1993: JP-346106.
PA (RIKA) RIKAGAKU KENKYUSHO.
DR WPI: 95-279918/37.
PT Oligo:nucleotide primer comprising amplification and sequencing
PT portions - useful for determination of fungal DNA sequences by PCR
PT amplification
PS Example 2: Page 6-7; 8pp; Japanese.
CC T02855-T02860 are amplification primers for DNA coding for
CC fungus-derived 18S rRNA. These primers may be bound at the 5' end
CC to the 3' end of a sequencing primer (T02861-T02863). The
CC resulting oligonucleotide primers comprising amplification and
CC sequencing portions (T02864-T02869). These primers are useful for
CC the determination of the base sequences of fungi.
SQ Sequence 39 BP; 12 A; 9 C; 9 G; 9 T;

```

```

Query Match 36.5%; Score 14.6; DB 1; Length 39;
Best Local Similarity 81.0%; Pred. No. 9e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 gacggccggcgctgatatca 21
DB 9 GACGGCCAGTCTGTCTCAA 29

RESULT 13
ID T36186/C
ID T36186 standard; DNA; 87 BP.
AC T36186;
DE 28-OCT-1996 (first entry)
DE Tagged protein DET1-DET2-spacer-SH2 DNA 5' primer.
KM Bone resorption disease; osteoporosis; src SH2 domain antagonist;
KM src homology 2 domain; polymerase chain reaction; PCR; primer; ss.
OS Synthetic.
PN EP-727211-A1.
PD 21-AUG-1996.
PR 07-FEB-1996; 200270.
PR 10-FEB-1995; US-386381.
PR 07-MAR-1995; US-400220.
PR 30-JUN-1995; US-497357.
PR 11-OCT-1995; US-541080.
PR 29-DEC-1995; US-580868.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PI Dunnington DJ;
DR WPI: 96-372674/38.
PT Use of selective src SH2 domain ligand - to prepare medicament for
PT treating bone resorption disease
PS Example 11; Page 17; 47pp; English.
CC A PCR primer pair (T36186 and T36187) was used to amplify a DNA
CC sequence encoding the tagged protein DET1-DET2-spacer-SH2, with
CC insertion of 5' NdeI and BamHI sites and 3' XbaI site, using a cDNA
CC clone contg. the chicken src gene as template. In the tagged
CC protein, DET1 is an epitope (W02128) from HIV-1 gp120 or gp160.
CC DET2 is a hexahistidine tag and SH2 is the chicken src SH2 domain.
CC DNA encoding DET1-DET2-spacer-SH2 was used to produce vectors
CC encoding fusion proteins (see also W02119-21) useful in binding
CC assays to determine the specificity of cpts. to inhibit SH2 domains;
CC cpts. that selectively inhibit the human src SH2 domain can be used
CC to treat bone resorption diseases such as osteoporosis.
SQ Sequence 87 BP; 20 A; 28 C; 20 G; 19 T;

Query Match 36.5%; Score 14.6; DB 1; Length 87;
Best Local Similarity 69.0%; Pred. No. 1e+03;
Matches 20; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 gacggccggcgctgatatca 29
DB 43 GACGGCCGGGCGACGCTGAATACGAATA 15

RESULT 14
ID T37293/C
ID T37293 standard; cDNA; 87 BP.
AC T37293;
DE 30-APR-1997 (first entry)
DE Chicken Src primer #1.
KM Polymerase chain reaction; PCR; amplify; primer; chicken; src;
KM SH2 domain; DET1; DET2; erythropoiesis; anaemia; haematopoiesis;
KM antagonist; ss.
OS Synthetic.
PN EP-728482-A2.
PD 28-AUG-1996.
PR 07-FEB-1996; 200269.
PR 10-FEB-1995; US-386381.
PR 07-MAR-1995; US-400220.
PR 30-JUN-1995; US-497357.
PR 11-OCT-1995; US-540680.

```

PR 29-DEC-1995: US-581089.
PA (SMK) SMITHKLINE BEECHAM CORP.
PI Dunnington DJ;
DR WPI: 96-386024/39.
PT Use of selective antagonist of haematopoietic acid phosphatase SH2
PT domain - with no significant affinity for other SH2 domains, to
PT increase erythropoiesis and haematopoiesis, esp. for treatment of
PT anaemia
PS Example 1: Page 17; 46pp; English.
CC The sequences given in T37293-94 are primers which were used in the
CC cloning of the chicken src SH2 domain containing tags DRT1 and DRT2.
CC The amplified sequence encodes a product which was used in the
CC isolation of a compound for improving erythropoiesis. The compound
CC may be used for the treatment of anaemia or to enhance haematopoiesis.
CC The isolated compound antagonises the hcp SH2 domain without side
CC effects caused by non-specific inhibition of other SH2 domains.
SQ Sequence 87 BP; 20 A; 28 C; 20 G; 19 T;

Query Match 36.5%; Score 14.6; DB 1; Length 87;
Best Local Similarity 69.0%; Pred. No. 1e+03;
Matches 20; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 gacggccggcgctgataatgacata 29
|||||
DB 43 GACGGCCGGCGCCACGCTGAATACGATA 15

RESULT 15
T51211/c
ID T51211 standard; DNA; 87 BP.
AC T51211;
DE 20-OCT-1997 (first entry)
DE Primer #1 for chicken src SH2 domain.
KW HIV-1; gp120; defined epitope tag; DRT1; envelope protein; human; Stat 6;
KW signal transduction and activation of transcription; src homology 2; PCR;
KW signaling molecule; protein tyrosine kinase; oncogenic protein; hcp SH2;
KW Grb2 SH2; allergic reaction; erythrocyte production; inhibitor; p85 SH2;
KW asthma; allergic rhinitis; atopic dermatitis; IGE receptor; SH-PTP2 SH2;
KW interleukin-4; IL-4; SH2 domain; Stat5 SH2; src SH2; lck SH2; fyn SH2;
KW IL-13; therapy; primer; amplify; polymerase chain reaction; ss.
OS Synthetic.
PN MO9702023-A1.
PD 23-JAN-1997.
PF 28-JUN-1996; U11074.
PF 08-FEB-1996; US-598716.
PR 30-JUN-1995; US-497357.
PA (SMK) SMITHKLINE BEECHAM CORP.
PI Dunnington DJ;
PI WPI: 97-108735/10.
PT Treating allergies with specific inhibitor of human Stat 6 SH2
PT domain - having very low binding affinity to panel of other SH2
PT domains so free of side effects, specifically for asthma and
PT allergic rhinitis
PS Example 11: Page 30; 86pp; English.
CC T51211-T51220 represent amplification primers used in the construction of
CC fusion proteins containing the src homology 2 (SH2) domain. The fusion
CC proteins are used to identify a compound that targets the human Stat
CC (signal transduction and activation of transcription) 6 SH2 domain. The
CC identified compounds have a binding affinity for Stat 6 over 50
CC (preferably 100) times higher than its affinity for the human Stat5 SH2
CC domain. The compound has an affinity for hcp SH2, SH-PTP2 SH2, p85 SH2,
CC Grb2 SH2, src SH2, lck SH2 or fyn SH2 of more than 50 (preferably 100)
CC times lower than its affinity for Stat 6 SH2. SH2 domains are conserved
CC non-catalytic sequences found in a variety of signalling molecules, such
CC as non-receptor protein tyrosine kinases, and in oncogenic proteins. The
CC compounds identified using the fusion proteins are used as the
CC administered compound in the method of the invention for treating
CC allergic reactions. Administration of the compound avoids the side
CC effects (e.g. reduced erythrocyte production) associated with
CC non-selective inhibition of SH2 domains. Selective compounds can be
CC identified in competitive binding assays using only a small subset (the
CC domains specified above) of SH2 domains rather than all 60 known domains.

CC The method can be used for the treatment of asthma and allergic rhinitis,
CC but can also be used to treat atopic dermatitis. Inhibition of the human
CC Stat 6 SH2 domain blocks up-regulation of the IGE receptor mediated by
CC interleukin-4 (IL-4) or IL-13.
SQ Sequence 87 BP; 20 A; 28 C; 20 G; 19 T;

Query Match 36.5%; Score 14.6; DB 1; Length 87;
Best Local Similarity 69.0%; Pred. No. 1e+03;
Matches 20; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 gacggccggcgctgataatgacata 29
|||||
DB 43 GACGGCCGGCGCCACGCTGAATACGATA 15

Search completed: June 4, 2000, 16:24:20
Job time: 28916 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:29 ; Search time 4521.53 Seconds
(without alignments)
35.857 Million cell updates/sec

Title: US-09-164-714-17

Perfect score: 40
Sequence: 1 gacggccggcgctgtatcaattgcatagcggttaagtt 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *
44: gb_est25: *

45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_est48: *
83: gb_est49: *
84: gb_est50: *
85: gb_est51: *
86: em_est35: *
87: em_est36: *
88: em_est37: *
89: em_est38: *
90: gb_est52: *
91: gb_est53: *
92: gb_est54: *
93: gb_est55: *
94: gb_est56: *
95: em_est39: *
96: em_est40: *
97: em_est41: *
98: em_est42: *
99: em_est43: *
100: em_est44: *
101: em_est45: *
102: gb_est57: *
103: gb_est58: *
104: em_est46: *
105: gb_est59: *
106: gb_est60: *
107: gb_est61: *
108: gb_est62: *
109: gb_est63: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	17.8	44.5	64	39	AA904844	AA904844 oJ80f12.s
2	17.6	44.0	56	70	AM156609	AM156609 se27e11.y
3	17	42.5	95	20	T52775	T52775 ya79g08.r1
4	16.6	41.5	73	37	AA683533	AA683533 zf34c12.s
5	16.2	40.5	56	42	AI079073	AI079073 SWBMLJ5BH
6	16	40.0	75	28	AA103683	AA103683 mo40b04.r
7	16	40.0	88	44	AI256675	AI256675 u130a03.y
8	16	40.0	90	37	AA688350	AA688350 nv15a05.s
9	15.8	39.5	79	21	T87474	T87474 y085e02.r1
10	15.6	39.0	60	44	AI318205	AI318205 ta52c01.x
11	15.6	39.0	99	41	AU011571	AU011571 AU011571
12	15.4	38.5	100	22	R58627	R58627 G4408 fetal
13	15.2	38.0	67	33	AA866073	AA866073 oh31h05.s
14	15.2	38.0	90	41	AU007806	AU007806 AU007806
15	15.2	38.0	90	41	AU007807	AU007807 AU007807
16	15.2	38.0	90	41	AF083179	AF083179 Capra_hir
17	15	37.5	88	39	AA864661	AA864661 oh37c11.s
18	15	37.5	90	34	AA503993	AA503993 nh39e01.s
19	15	37.5	94	43	AI172856	AI172856 uc10c09.r
20	15	37.5	94	62	AI936176	AI936176 wo62q09.x
21	15	37.5	97	41	AI034469	AI034469 ow24d12.x
22	15	37.5	98	23	H39547	H39547 y168h07.r1
23	14.8	37.0	65	38	AA755665	AA755665 vs74e05.r
24	14.8	37.0	77	34	AA494110	AA494110 ne54b01.s
25	14.8	37.0	77	39	AA839381	AA839381 vw48h02.r
26	14.6	36.5	49	33	AA433110	AA433110 v094g03.r
27	14.6	36.5	67	47	AI491935	AI491935 to06g01.x
28	14.6	36.5	69	21	R06240	R06240 yeg94h09.r1
29	14.6	36.5	74	45	AJ237129	AJ237129 AJ237129
30	14.6	36.5	100	40	AA928885	AA928885 ya44d11.r
31	14.4	36.0	52	20	D25845	D25845 HUMGS04217
32	14.4	36.0	58	27	C00161	C00161 HUMGS000581
33	14.4	36.0	61	28	AA091692	AA091692 ml168.seq
34	14.4	36.0	64	37	AA716059	AA716059 zh32f12.s
35	14.4	36.0	73	39	AA878250	AA878250 o104f02.s
36	14.4	36.0	76	25	N89340	N89340 K9244F Huma
37	14.4	36.0	79	44	AI272143	AI272143 ap61e04.x
38	14.4	36.0	81	23	HA2950	HA2950 y068c07.r1
39	14.4	36.0	81	23	R86045	R86045 y062a01.r1
40	14.4	36.0	82	36	AA664777	AA664777 nu68f02.s
41	14.4	36.0	82	62	AI906839	AI906839 RC-BT126-
42	14.4	36.0	82	62	AI924362	AI924362 wn55h10.x
43	14.4	36.0	88	62	AI902577	AI902577 QV-BT009-
44	14.4	36.0	91	21	T81025	T81025 yd25d10.s1
45	14.4	36.0	91	21	T81025	T81025 yd25d10.s1

ALIGNMENTS

RESULT 1

AA904844 64 bp mRNA EST 09-JUN-1998
 LOCUS OJ80f12.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
 DEFINITION IMAGE:1504655 3' similar to TR:000405 000405 FB19 PROTEIN. ;, mRNA
 sequence.

AA904844
 VERSION AA904844.1 GI:3039967
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 64)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov

This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Trace considered overall poor quality
 Insert length: 383 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers
 1. 64
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1504655"
 /clone_lib="Soares_NFL_T_GBC_S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
 a modified polylinker; Site: 1: Not I; Site: 2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NHH19W, testis NHR, and B-cell
 NCI-CCAP-GCB1) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonalido."

BASE COUNT 9 a 14 c 31 g 10 t
 ORIGIN

Query Match 44.5%; Score 17.8; DB 39; Length 64;
 Best Local Similarity 67.6%; Pred. No. 6; le=02;
 Matches 25; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Oy 4 gggccggcgtgtatcatatgcatagcggtgaatt 40
 Db 10 ggcctggcggcggatattgtggccggcggaatt 46

RESULT 2

AM156609 56 bp mRNA EST 04-NOV-1999
 LOCUS se27e11.y1 Gm-c1015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 DEFINITION Gm-c1015-2469 5', mRNA sequence.

ACCESSION AM156609
 VERSION AM156609.1 GI:6227947
 KEYWORDS EST.
 SOURCE soybean.

ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
 Glycine.

REFERENCE 1 (bases 1 to 56)
 Shoemaker R., Kelm P., Vodkin L., Epeiding J., Coryell V.,
 Khanna A., Bolla B., Marra M., Hillier L., Kucaba T., Martin J.,
 Beck C., Wylie T., Underwood K., Steptoe M., Thelshing B., Allen M.,
 Bowers Y., Person B., Swaller T., Gibbons M., Pape D., Harvey N.,
 Schurk R., Ritter E., Kohn S., Shih T., Jackson Y., Cardenas M.,
 McCann R., Waterston R. and Wilson R.

JOURNAL Public Soybean EST Project
 COMMENT On Jul 8, 1999 this sequence version replaced gi:5421748.
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134 For further information
 call: (800) 430-0030 or (314) 427-3322 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or

info@genomesystems.com web site: www.genomesystems.com
Seq primer: -40RP from gibco.

FEATURES
source

1. .56
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1015-2469"
/clone_lib="Gm-c1015"
/tissue_type="Mature flowers, field grown plants"
/lab_host="Xl10-Gold"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from mature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into Xl10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Expelding."

BASE COUNT 13 a 4 c 16 g 23 t
ORIGIN

Query Match 44.0%; Score 17.6; DB 70; Length 56;
Best Local Similarity 71.9%; Pred. No. 7.2e+02;
Matches 23; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9 gggctgtatcaatggcataagcgtaagtt 40
| | | | | | | | | | | | | | | | | | | |
Db 23 GAGTAGTTAAATGCGACGTGAGCTAATT 54

RESULT 3

LOCUS T52775 95 bp mRNA EST 06-FEB-1995
DEFINITION ya79d08.r1 Stratagene ovary (#937217) Homo sapiens cDNA clone IMAGE:67887 5' similar to contains Alu repetitive element, mRNA sequence.

ACCESSION T52775
VERSION T52775.1 GI:654635

KEYWORDS
SOURCE

ORGANISM human.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 95)

ILLIER,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chascoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Madis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.

Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

TITLE JOURNAL MEDLINE
COMMENT 97044478
Other ESTs: ya79d08.s1

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Insert Size: 1013
High quality sequence stops: 82 Source: IMAGE Consortium, LNL. This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1013 Std Error: 0.00

Seq primer: M13Rp1
High quality sequence stop: 82.

FEATURES
source

1. .95
/organism="Homo sapiens"
/db_xref="GDB:504144"
/db_xref="taxon:9606"
/clone="IMAGE:67887"
/clone_lib="Stratagene ovary (#937217)"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ovary; Vector: Bluescript SK; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Total ovary tissue, normal, caucasian. Average insert size: 0.8 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GATTCGACGACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTT TTTT TTTT TTTT TTTT 3' "

BASE COUNT 19 a 32 c 20 g 24 t
ORIGIN

Query Match 42.5%; Score 17; DB 20; Length 95;
Best Local Similarity 80.0%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 gcccggtgtatcaatggcata 29
| | | | | | | | | | | | | | | | | | | |
Db 12 GCCCAGCGCTGTACAGTGCAC 36

RESULT 4
LOCUS AA683533/c 73 bp mRNA EST 19-DEC-1997
DEFINITION zf34c12.s1 Soares fetal heart NBH19W Homo sapiens cDNA clone IMAGE:378838 3' similar to gb:M24194 GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT-LIKE PROTEIN (HUMAN);, mRNA sequence.

ACCESSION AA683533
VERSION AA683533.1 GI:26710131

KEYWORDS
SOURCE

ORGANISM human.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 73)

ILLIER,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,T., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R. WashU-NCI human EST Project
Unpublished (1997)

On Sep 12, 1996 this sequence version replaced gi:1394601.
Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amerisham.

FEATURES
source

1. .73
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:378838"
/clone_lib="Soares_fetal_heart_NBH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: heart; Vector: pRT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - Oligo(dT) primer [5'

TGTACCAATCTGAAGTGGAGCGCCGCATCTTTTCTTTTCTTTT 3',
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library went through one round of
 normalization to a cot = 5. Library constructed by
 M.Felina Bonaldo. This library was constructed from the
 same fetus as the fetal lung library, Soares fetal lung
 NBHL19W."

BASE COUNT 18 a 17 c 13 g 25 t
 ORIGIN

Query Match 41.5%; Score 16.6; DB 37; Length 73;
 Best Local Similarity 71.0%; Pred. No. 2.1e+03;
 Matches 22; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 10 ggcgtgtatcaatggcgaagcgttaagt 40
 ||| ||| | ||||| | ||| |||
 Db 73 GGCAGGTGACATGTGCACGCTAGAAGTT 43

RESULT 5
 A1079073 56 bp mRNA EST 12-AUG-1998
 LOCUS SWBML3SBH11G10SK Brugia malayi L3 subtracted cDNA library
 DEFINITION (SAM97YG-Bml3SB) Brugia malayi cDNA clone SWBML3SBH11G10 5' or 3',
 mRNA sequence.
 ACCESSION A1079073
 VERSION A1079073.1 GI:3415324
 KEYWORDS EST.
 SOURCE Brugia malayi.
 ORGANISM Eukaryota; Metazoa; Nematoda; Secernentea; Spirurida; Spirurida;
 Filarioidea; Onchocercidae; Brugia.
 REFERENCE 1 (bases 1 to 56)
 AUTHORS Williams, S.A.
 TITLE Genes expressed in infective third stage larvae of Brugia malayi
 JOURNAL Unpublished (1995)
 COMMENT On May 8, 1995 this sequence version replaced gi:801111.
 CONTACT: Steven A. Williams
 MOLECULAR PARASITOLOGY
 Smith College Department of Biological Sciences
 Department of Biological Sciences, Clark Science Center, Smith
 College, Northampton, MA, 01063, USA
 Tel: 4135853826
 Fax: 4135853786
 Email: genome@smith.edu
 Seq primer: bluescript SK.
 FEATURES
 source
 1..56
 Location/Qualifiers
 /organism="Brugia malayi"
 /db_xref="taxon:6279"
 /clone="SWBML3SBH11G10"
 /clone_1lb="Brugia malayi L3 subtracted cDNA library
 (SAM97YG-Bml3SB)"
 /dev_stage="conventional L3 cDNA minus spliced-leader L3
 cDNA"
 /lab_host="XLI-Blue MRF"
 /note="Vector: Lambda ZAP II (stratagene); Site_1: Eco RI;
 Site_2: Eco RI; Lymphatic filarial nematode parasite of
 humans. cDNA inserts were obtained by subtracting the
 JH093SL-Bml3 cDNA library (constructed using SL sequence
 and PCR) from SAM94WL-Bml3 cDNA library (constructed
 using conventional cDNA methods). cDNA was amplified from
 both libraries using PCR and SL cDNA clones were
 eliminated through hybridization. The library has 5 x
 10E6 independent recombinants and the average insert size
 is ~ 600 bp. The library was constructed by Ying Gong.
 The library is available from Dr.S.A.Williams, email:
 genome@smith.edu."

BASE COUNT 16 a 12 c 11 g 17 t
 ORIGIN

Query Match 40.5%; Score 16.2; DB 42; Length 56;
 Best Local Similarity 72.4%; Pred. No. 2.9e+03;
 Matches 21; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 gacggccggcgtgtatcaatggcata 29
 | | ||||| ||| | | | | |
 Db 7 GGCAGGTGACATGTGCACGCTAGAAGTT 35

RESULT 6
 A103683 75 bp mRNA EST 29-OCT-1996
 LOCUS m40b04.r1 life Tech mouse embryo 15 5dpc 10667012 Mus musculus
 DEFINITION cDNA clone IMAGE:556015 5', mRNA sequence.
 ACCESSION A103683
 VERSION A103683.1 GI:1649846
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 75)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMNI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Nov 29, 1993 this sequence version replaced gi:430148.
 CONTACT: Marra M/Mouse EST Project
 WashU-HMNI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LBNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:336807
 Seq primer: -28M13 rev1 from Amersham
 High quality sequence stop: 72.
 FEATURES
 source
 1..75
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:556015"
 /clone_1lb="Life Tech mouse embryo 15 5dpc 10667012"
 /tissue_type="embryo"
 /dev_stage="15.5dpc embryos"
 /lab_host="DH10B"
 /note="Organ: whole embryo; Vector: PCMV-SPORT2; Site_1:
 SalI; Site_2: NotI; Cloned unidirectionally. Primer:
 Oligo dT. 15.5dpc embryos. PCMV-SPORT2 vector."

BASE COUNT 20 a 16 c 19 g 20 t
 ORIGIN

Query Match 40.0%; Score 16; DB 28; Length 75;
 Best Local Similarity 79.2%; Pred. No. 3.8e+03;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 ggcgtgtatcaatggcgaagc 32
 ||||| ||| | | | | |
 Db 17 GGCAGGTGACATGTGCACGCTAGAAGTT 40

RESULT 7
 A1256675/c 88 bp mRNA EST 12-NOV-1998
 LOCUS A1256675

DEFINITION u130a03.y1 Soares mouse urogenital ridge NMUR Mus musculus cDNA
1 clone IMAGE:1852780 5', mRNA sequence.
ACCESSION A1256675
VERSION A1256675.1 GI:3864200
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 88)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, R., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HM Mouse EST Project
Unpublished (1996)
On Jan 19, 1998 this sequence version replaced gi:2284957.
CONTACT: Marras M/Mouse EST Project
WashU-HM Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:968208
Seq primer: -40RP from Gibco
High quality sequence stop: 58.
Location/Qualifiers
FEATURES
Source
1..88
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1852780"
/clone_lib="Soares mouse urogenital ridge NMUR"
/sex="equal ratio of male:female"
/tissue_type="urogenital ridge (embryonic)"
/dev_stage="fetal, mixture of 11.5 and 12.5 dpc"
/lab_host="DH10B"
/note="Organ: gonad; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAGAGTGGAGGCGCCGCAATTCCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library went through two rounds of normalization, and was
constructed by Bento Soares and M.Fatima Bernaldo."

BASE COUNT 21 a 25 c 18 g 24 t
ORIGIN

Query Match 40.0%; Score 16; DB 44; Length 88;
Best Local Similarity 62.5%; Pred. NO. 4e+03;
Matches 25; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 gacgccccggctcgatcaatgcatagcgagtaagt 40
| | | | | | | | | | | | | | | | | | | | | |
Db 77 GCCTGTCGAGATGCTGCTGCGATCTGCTGTGAGCT 38

RESULT 8
LOCUS AA688350 90 bp mRNA EST 11-DEC-1997
DEFINITION nv15a05.s1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:1220240 3',
mRNA sequence.
ACCESSION AA688350
VERSION AA688350.1 GI:2675256
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

REFERENCE 1 (bases 1 to 90)
Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Nov 6, 1997 this sequence version replaced gi:315230.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.llnl.gov/bdrp/image/image.html
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 78.
Location/Qualifiers
FEATURES
Source
1..90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1220240"
/clone_lib="NCI_CGAP_Pr22"
/sex="male"
/tissue_type="normal prostate"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; 1st strand cDNA was prepared
from normal prostate bulk tissue, and was then primed with
a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized, and was
constructed by Bento Soares and M. Fatima Bernaldo."

BASE COUNT 36 a 6 c 17 g 31 t
ORIGIN

Query Match 40.0%; Score 16; DB 37; Length 90;
Best Local Similarity 79.2%; Pred. NO. 4e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 ggtatcaattgcatagcgagta 37
| | | | | | | | | | | | | | | | | | | | | |
Db 48 GCTATCATTTGACACAGGGGGA 71

RESULT 9
LOCUS T87474/C 79 bp mRNA EST 17-MAR-1995
DEFINITION y85e02.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
IMAGE:115034 5' similar to SP:S27875 S27875 GENE PC326 PROTEIN - ;,
mRNA sequence.
ACCESSION T87474
VERSION T87474.1 GI:715826
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 79)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marras, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK

FEATURES
source National Institute of Radiological Sciences
 9-1, Anagawa-4-chome, Inage-ku, Chiba 263-8555, Japan
 Email: mortimyo@nirs.90.jp.
 Location/Qualifiers
 1.90

```

source
1. .90
/organism="Schizosaccharomyces pombe"
/strain="972"
/db_xref="taxon:4896"
/clone="spc02527"
/clone_lib="Schizosaccharomyces pombe late log phase cDNA"
/sex="h minus"
/note="vector: M13mp19; The cDNA library of
Schizosaccharomyces pombe was prepared by cloning cDNA
into the SmaI site of M13mp19 DNA and the direction of DNA
sequences was not always from 5' to 3'. The cDNA data of
Schizosaccharomyces pombe are available for searching on
the World Wide Web. (URL, http://www.nhrs.go.jp)"
24 a 29 c 16 g 21 t
BASE COUNT
ORIGIN

```

Query Match	38.0%;	Score 15.2;	DB 41;	Length 90;
Best Local Similarity	71.4%;	Pred. No. 8.9e+03;		
Matches 20; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;

OY 10 ggcctgtatcaattgcatagcggtaa 37
||| ||| ||| ||| ||
Db 68 GGCCTGTTCTCGCTGGCAACGCCGAAA 41

RESULT	15			
AU007807/c				
LOCUS	AU007807	90 bp	mRNA	EST
DEFINITION	AU007807 Schizosaccharomyces pombe late log phase cDNA			31-JUL-1996
ACCESSION	Schizosaccharomyces pombe cDNA clone spc02528,			mRNA sequence.
VERSION	AU007807			
KEYWORDS	AU007807.1 GI:3344265			
SOURCE	EST.			
ORGANISM	Fission yeast.			
	Schizosaccharomyces pombe			

REFERENCE	1 (bases 1 to 90)
AUTHORS	Moriya,M. and Mita,K.
TITLE	Identification of expressed sequence tags of Schizosaccharomyces pombe
JOURNAL	unpublished (1998)
COMMENT	On May 9, 1995 this sequence version replaced gi:803109.

JOURNAL Unpublished (1998)
 COMMENT On May 9, 1995 this sequence version replaced gi:803109.
 CONTACT Mitsuki Moriyono
 GENOME RESEARCH GROUP
 NATIONAL INSTITUTE OF RADIOLOGICAL SCIENCES
 9-1, Aragawa-4-chome, Inage-ku, Chiba 263-8555, Japan
 EMAIL: moriyono@ris.ri.go.jp
 LOCATION/QUALIFIERS
 SOURCE 1. 90

FEATURES	SOURCE
Location/Qualifiers	1. . 90
	/organism="Schizosaccharomyces pombe"
	/strain="972"
	/db.xref="taxon:4896"
	/clone="spc02528"
	/clone_id="Schizosaccharomyces pombe late log phase cDNA"
	/sex="h minus"
	/note="Vector: M13mp19: The cDNA library of
	Schizosaccharomyces pombe was prepared by cloning cDNA
	into the SmaI site of M13mp19 DNA and the direction of DNA
	sequences was not always from 5' to 3'. The cDNA data of
	Schizosaccharomyces pombe are available for searching on
	the World Wide Web. (URL, http://www.nlrs.go.jp)"
BASE COUNT	24 a 29 c 16 g 21 t
ORIGIN	

Query Match	38.0%;	Score 15.2;	DB 41;	Length 90;
Best Local Similarity	71.4%;	Pred. No. 8.9e+03;		

	Matches	20;	Conservative	0;	Mismatches	8;	Indels	0;	Gaps	0
QY	10	gycgtgtatcaat	tygcacatagcg	cgtaa	37					
Db	68	ggcttggttttct	ctctggcaacggc	ggaaa	41					

```
Search completed: June  4, 2000, 13:53:32
Job time: 20956 sec
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:46 ; Search time 244.64 Seconds
(without alignments)
21.253 Million cell updates/sec

Title: US-09-164-714-17

Perfect score: 40

Sequence: 1 gacggccggcgctgtatcaattgcataggcgtaagt 40

Scoring table: IDENTITY_NUC

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
7: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	15.8	39.5	38	5	US-08-718-904-124
2	15.6	39.0	51	3	US-08-432-871C-25
3	15	37.5	94	2	US-08-472-255A-24
4	15	37.5	94	2	US-08-479-724A-24
5	15	37.5	94	2	US-08-472-256B-24
6	15	37.5	94	6	PCT-US96-09455A-24
7	14.6	36.5	35	2	US-08-446-856A-2
8	14.4	36.0	93	2	US-08-388-672A-11
9	14.4	36.0	93	5	US-09-080-554-11
10	14.4	36.0	97	6	PCT-US96-09455A-284
11	14.2	35.5	49	1	US-08-298-073-22
12	14.2	35.5	49	2	US-08-794-153-22
13	14.2	35.5	73	2	US-08-434-001-64
14	14.2	35.5	73	2	US-08-433-585-64
15	14.2	35.5	73	2	US-08-434-425-64
16	14.2	35.5	73	3	US-08-437-667-64
17	14.2	35.5	73	5	US-08-906-955-64
18	14.2	35.5	73	6	PCT-US96-06060-64
19	14	35.0	22	2	US-08-468-580-20
20	14	35.0	22	6	PCT-US95-03731-20
21	14	35.0	31	1	US-08-457-245-18
22	14	35.0	37	3	US-08-880-557-35
23	14	35.0	63	1	US-08-343-943-13
24	14	35.0	96	6	PCT-US96-09455A-274
25	13.8	34.5	23	1	US-08-010-099-15
26	13.8	34.5	23	1	US-08-010-099-15
27	13.8	34.5	23	1	US-08-010-099-15

c 28	13.8	34.5	23	1	US-08-010-099-22	Sequence 22, Appl
c 29	13.8	34.5	23	2	US-08-448-716-6	Sequence 6, Appl
c 30	13.8	34.5	23	2	US-08-448-716-15	Sequence 15, Appl
c 31	13.8	34.5	23	2	US-08-448-716-18	Sequence 18, Appl
c 32	13.8	34.5	23	2	US-08-448-716-22	Sequence 22, Appl
c 33	13.8	34.5	55	6	PCT-US96-09451-71	Sequence 71, Appl
c 34	13.6	34.5	77	1	US-08-399-412A-14	Sequence 14, Appl
c 35	13.6	34.0	48	1	US-08-140-797-2	Sequence 2, Appl
c 36	13.6	34.0	48	1	US-08-486-670A-2	Sequence 2, Appl
c 37	13.6	34.0	66	2	US-08-323-531-34	Sequence 34, Appl
c 38	13.6	34.0	66	2	US-08-198-094-34	Sequence 34, Appl
c 39	13.6	34.0	66	6	PCT-US93-07424-34	Sequence 34, Appl
c 40	13.6	34.0	66	6	PCT-US95-02087-34	Sequence 34, Appl
c 41	13.6	34.0	78	1	US-08-411-796-175	Sequence 175, App
c 42	13.6	34.0	78	5	US-08-471-039-175	Sequence 175, App
c 43	13.6	34.0	78	6	PCT-US93-11198-175	Sequence 175, App
c 44	13.4	33.5	43	2	US-08-292-081A-13	Sequence 13, Appl
c 45	13.4	33.5	43	3	US-08-752-495-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-718-904-124
Sequence 124, Application US/08718904
Patent No. 6037329
GENERAL INFORMATION:
APPLICANT: Baird, J. Andrew
APPLICANT: Chandler, Lois Ann
APPLICANT: Sosnowski, Barbara A.
TITLE OF INVENTION: COMPOSITIONS CONTAINING NUCLEIC ACIDS AND LIGANDS FOR THERA
NUMBER OF SEQUENCES: 128
CORRESPONDENCE ADDRESSES:
ADDRESS: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,904
FILING DATE: 24-SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: NO. 6037329tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 760100.415C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 124:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-718-904-124

Query Match 39.5%; Score 15.8; DB 5; Length 38;
Best Local Similarity 74.1%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
DB 8 TGTCTCATTTCCCTTGACCGGTAACT 34

RESULT 2
US-08-432-871C-25
; Sequence 25, Application US/08432871C
; Patent No. 5877010
; GENERAL INFORMATION:
; APPLICANT: loeb, Lawrence A.
; APPLICANT: Black, Margaret E.
; TITLE OF INVENTION: THYMIDINE KINASE MUTANTS
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432.871C
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 240052.409C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-432-871C-25

Query Match 39.0%; Score 15.6; DB 3; Length 51;
Best Local Similarity 53.8%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 2 acgcccggcggtatcataatgcatagcggttaagt 40
DB 8 ACCGCGCGGCGCGGTAGCANNNNNNNNGCGATGGGAT 46

RESULT 3
US-08-472-255A-24
; Sequence 24, Application US/08472255A
; Patent No. 5766853
; GENERAL INFORMATION:
; APPLICANT: PARMA, DAVID
; APPLICANT: GOLD, LARRY
; TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS
; TITLE OF INVENTION: TO SELECTINS (AS AMENDED)
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS

SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,255A
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX40-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 94 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All C's are 2'-NH2 cytosine
; FEATURE:
; OTHER INFORMATION: All U's are 2'-NH2 uracil
; US-08-472-255A-24

Query Match 37.5%; Score 15; DB 2; Length 94;
Best Local Similarity 56.4%; Pred. No. 2.7e+02;
Matches 22; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

QY 2 acgcccggcggtatcataatgcatagcggttaagt 40
DB 35 AUGCGCGGGGCAAGAGUGGUGGCGCAACGAGCGCCAGCU 73

RESULT 4
US-08-479-724A-24
; Sequence 24, Application US/08479724A
; Patent No. 5780228
; GENERAL INFORMATION:
; APPLICANT: PARMA, DAVID
; APPLICANT: GOLD, LARRY
; TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS
; TITLE OF INVENTION: TO LECTINS
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,724A
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991


```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX40-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ. ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 94 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All C's are 2'-NH2 cytosine
; FEATURE:
; OTHER INFORMATION: All U's are 2'-NH2 uracil
US-08-479-724A-24

Query Match          37.5%; Score 15; DB 2; Length 94;
Best Local Similarity 56.4%; Pred. No. 2.7e+02;
Matches 22; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

OY 2 acggccgggctgtatcaattggcatagcggtgaattc 40
      ||||| ||||| ||||| ||||| ||||| |||||
DB 35 AUGCCGGGGCAGGAGAGUGGGUGGCAACGAGGCCAGCU 73

RESULT 5
US-08-472-256B-24
; Sequence 24, Application US/08472256B
; Patent No. 6001988
; GENERAL INFORMATION:
; APPLICANT: PARMA, DAVID
; APPLICANT: GOLD, LARRY
; TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID LIGANDS
; TITLE OF INVENTION: TO LECTINS
; NUMBER OF SEQUENCES: 177
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,256B
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson

```

```

; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ. ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 94 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All C's are 2'-NH2 cytosine
; FEATURE:
; OTHER INFORMATION: All U's are 2'-NH2 uracil
US-08-472-256B-24

Query Match          37.5%; Score 15; DB 5; Length 94;
Best Local Similarity 56.4%; Pred. No. 2.7e+02;
Matches 22; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

OY 2 acggccgggctgtatcaattggcatagcggtgaattc 40
      ||||| ||||| ||||| ||||| ||||| |||||
DB 35 AUGCCGGGGCAGGAGAGUGGGUGGCAACGAGGCCAGCU 73

RESULT 6
PCT-US96-09455A-24
; Sequence 24, Application PC/TUS9609455A
; GENERAL INFORMATION:
; APPLICANT: PARMA, et al.
; TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID
; TITLE OF INVENTION: LIGANDS TO LECTINS
; NUMBER OF SEQUENCES: 390
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/09455A
; FILING DATE: 05 JUNE 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,724
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/472,256
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/472,255
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/477,829
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX40C/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ. ID NO: 24:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 94 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All C's are 2'-NH2 cytosine
FEATURE:
OTHER INFORMATION: All U's are 2'-NH2 uracil
PCT-US96-09455A-24

Query Match 37.5%; Score 15; DB 6; Length 94;
Best Local Similarity 56.4%; Pred. No. 2.7e+02;
Matches 22; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

QY 2 acgcccggcgtgtatcataatggcgtagcggtatgct 40
| | | | | | | | | | | | | | | | | | | | | |
Db 35 AUGCCCGGGCAGAGAGUGGGGCGACGAGCGCCAGCU 73

RESULT 7
US-08-446-856A-2
; Sequence 2, Application US/08446856A
; Patent No. 5776675
; GENERAL INFORMATION:
; APPLICANT: Broad, Peter Michael
; TITLE OF INVENTION: Identification of Compounds
; TITLE OF INVENTION: Modulating Protein/Cell Membrane Interaction
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,856A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/02543
; FILING DATE: 14-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9226065.2
; FILING DATE: 14-DEC-1992
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-446-856A-2

Query Match 36.5%; Score 14.6; DB 2; Length 35;
Best Local Similarity 81.0%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ggcgcggcgtgtatcataatggcgtatgct 24
| | | | | | | | | | | | | | | | | | | | | |
Db 7 GGCCCGCCTGTGATCAATG 27

RESULT 8
US-08-388-672A-11
; Sequence 11, Application US/08388672A

; Patent No. 5795961
; GENERAL INFORMATION:
; APPLICANT: Wallace, T. Paul
; APPLICANT: Harris, William J.
; APPLICANT: Carr, Frank J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Kilmura, Kunio
; APPLICANT: Welt, Sydney
; TITLE OF INVENTION: Recombinant Human Anti-Lewis B
; TITLE OF INVENTION: Antibodies
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte and Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,672A
FILING DATE: 14-FEB-1995

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5795961man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-388-672A-11

Query Match 36.0%; Score 14.4; DB 2; Length 93;
Best Local Similarity 65.6%; Pred. No. 5e+02; 11;
Matches 21; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 4 ggcgcggcgtgtatcataatggcgtatgct 35
| | | | | | | | | | | | | | | | | | | | | |
Db 4 GGCTTCTGCTGGTACCAATATTAATAGTGTT 35

RESULT 9
US-09-080-554-11
; Sequence 11, Application US/09080554
; Patent No. 6025481
; GENERAL INFORMATION:
; APPLICANT: Wallace, T. Paul
; APPLICANT: Harris, William J.
; APPLICANT: Carr, Frank J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Welt, Sydney
; APPLICANT: Kilmura, Kunio
; TITLE OF INVENTION: Recombinant Human Anti-Lewis B
; TITLE OF INVENTION: Antibodies
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.

ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/080,554
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/388,672
FILING DATE: 14-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: 7606-032
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-09-080-554-11

Query Match 36.0%; Score 14.4; DB 5; Length 93;
Best Local Similarity 65.6%; Pred. No. 5e+02;
Matches 21; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Db 4 GGCTTCTGCTGTAACATATTAATAGTGTT 35

RESULT 10
PCT-US96-09455A-284
Sequence 284, Application PC/TUS9609455A
GENERAL INFORMATION:
APPLICANT: PARMA, et al.
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC ACID
TITLE OF INVENTION: LIGANDS TO LECTINS
NUMBER OF SEQUENCES: 390
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09455A
FILING DATE: 05 JUNE 1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,724
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA: 08/472,256
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA: 08/472,255
APPLICATION NUMBER: 08/472,255
FILING DATE: 07-JUNE-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,829
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX40C/PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 284:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All C's are 2'-NH2 cytosine
OTHER INFORMATION: All U's are 2'-NH2 uracil
PCT-US96-09455A-284

Query Match 36.0%; Score 14.4; DB 6; Length 97;
Best Local Similarity 56.2%; Pred. No. 5.1e+02;
Matches 18; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

Db 22 AAGUACCGGAGGAGGACUGGAGUAGGCG 53

Query 11
US-08-298-073-22
Sequence 22, Application US/08298073
Patent No. 5681702
GENERAL INFORMATION:
APPLICANT: COLLINS, MARK L.
APPLICANT: HORN, THOMAS
APPLICANT: SHERIDAN, PATRICK E.
APPLICANT: WARNER, BRIAN D.
APPLICANT: URDEA, MICHAEL S.
TITLE OF INVENTION: REDUCTION OF NONSPECIFIC HYBRIDIZATION
TITLE OF INVENTION: BY USING NOVEL, BASE-PAIRING SCHEMES
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION, INTELLECTUAL PROPERTY -
ADDRESS: R440
STREET: P.O. BOX 8097
CITY: EMERYVILLE
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/298,073
FILING DATE: 30-AUG-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: GOLDMAN, KENNETH M.
REGISTRATION NUMBER: 34,174
REFERENCE/DOCKET NUMBER: 0974,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2719
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 base pairs


```

APPLICANT: CHEN, HANG
APPLICANT: MORRIS, KEVIN
APPLICANT: STEPHENS, ANDREW
APPLICANT: GOLD, LARRY
TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS BY
TITLE OF INVENTION: EXPONENTIAL ENRICHMENT: TISSEX
TITLE OF INVENTION: SELEX
NUMBER OF SEQUENCES: 235
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
COMPUTER: IBM pc compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/433,585
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX30.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 73 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-433-585-64

Query Match 35.5%; Score 14.2; DB 2; Length 73;
Best Local Similarity 62.9%; Pred. No. 5,9e+02;
Matches 22; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

07 4 ggcgcggcgtggtatcattggcgcataagcggtgaag 38
      |||| | | |||| | |||| | |||| | ||||
Db 67 ggcctcatgtcgatcaccgccggatagccggagag 33

RESULT 15
US-08-434-425-64/C
; Sequence 64, Application US/08434425
; Patent No. 5789157
; GENERAL INFORMATION:
; APPLICANT: JENSEN, KIRK
; APPLICANT: CHEN, HANG
; APPLICANT: MORRIS, KEVIN
; APPLICANT: STEPHENS, ANDREW
; APPLICANT: GOLD, LARRY
; TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS BY
; TITLE OF INVENTION: EXPONENTIAL ENRICHMENT: TISSEX
; TITLE OF INVENTION: SELEX
; NUMBER OF SEQUENCES: 235
; CORRESPONDENCE ADDRESSES:

```

```

      ADDRESSSEE: Swanson & Bratschun, L.L.C.
      STREET: 8400 E. Prentice Avenue, Suite 200
      CITY: Englewood
      STATE: Colorado
      COUNTRY: USA
      ZIP: 80111

COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
  COMPUTER: IBM pc compatible
  OPERATING SYSTEM: MS-DOS
  SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/434,425
  FILING DATE:
  CLASSIFICATION: 536
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/714,131
  FILING DATE: 10-JUNE-1991
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/536,428
  FILING DATE: 11-JUNE-1990
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/964,624
  FILING DATE: 21-OCTOBER-1992
  ATTORNEY/AGENT INFORMATION:
  NAME: Barry J. Swanson
  REGISTRATION NUMBER: 33,215
  REFERENCE/DOCKET NUMBER: NEX30.1
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: (303) 793-3333
  TELEFAX: (303) 793-3433
  INFORMATION FOR SEQ ID NO: 64:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 73 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear

US-08-434-425-64

Query Match      35.5%; Score 14.2; DB 2; Length 73;
Best Local Similarity 62.9%; Pred. No. 5.9e+02;
Matches 22; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

Search completed: June 4, 2000, 16:09:47
Job time: 28066 sec

QY 4 ggcgcggtctgtatcaattgcatggcgtaag 38
||| | | ||| | | ||| | | |||
Db 67 ggccatgttcgaatcaccgggatgccgggaag 33

QY 4 ggccgggctgtatcaatcattgcatagcgctaa 38
||| | |||| | |||| |||
Db 67 ggccatgtcgaatcacccgggaatgccggaag 33

RESULT 15
 US-08-434-425-64/C
 : Sequence 64, Application US/08434425
 : Patent No. 5789157
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: JENSEN, KIRK
 : APPLICANT: CHEN, HANG
 :
 : APPLICANT: MORRIS, KEVIN
 : APPLICANT: STEPHENS, ANDREW
 :
 : APPLICANT: GOLD, LARRY
 :
 : TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS BY
 : TITLE OF INVENTION: EXPONENTIAL ENRICHMENT: TISSUE
 : TITLE OF INVENTION: SELEX
 :
 : NUMBER OF SEQUENCES: 235
 :
 : CORRESPONDENCE ADDRESS:

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:43 ; Search time 1236.38 Seconds

(Without alignments)
-29.899 Million cell updates/sec

Title: US-09-164-714-18

Perfect score: 38

Sequence: 1 catgctgcagcttgaccacatttggcaccacaataggg 38

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 882769 segs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pl4:*
11: gb_pl5:*
12: gb_pl6:*
13: gb_pl7:*
14: gb_pl8:*
15: gb_pl9:*
16: gb_pl10:*
17: gb_pl11:*
18: gb_pl12:*
19: gb_pl13:*
20: gb_pl14:*
21: gb_pl15:*
22: gb_pl16:*
23: gb_pl17:*
24: gb_pl18:*
25: gb_pl19:*
26: gb_pl20:*
27: gb_pl21:*
28: gb_pl22:*
29: gb_pl23:*
30: gb_pl24:*
31: gb_pl25:*
32: gb_pl26:*
33: gb_pl27:*
34: gb_pl28:*
35: gb_pl29:*
36: gb_pl30:*
37: gb_pl31:*
38: gb_pl32:*
39: gb_pl33:*
40: gb_pl34:*
41: gb_pl35:*
42: gb_pl36:*
43: gb_pl37:*
44: gb_pl38:*

45: gb_pl39:*
46: gb_pl40:*
47: gb_pl41:*
48: gb_pl42:*
49: gb_pl43:*
50: gb_pl44:*
51: gb_pl45:*
52: gb_pl46:*
53: gb_pl47:*
54: gb_pl48:*
55: gb_pl49:*
56: gb_pl50:*
57: gb_pl51:*
58: gb_pl52:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	17.4	45.8	60	1	S77980	S77980 comp1-major
2	17.4	45.8	60	2	AF065436	AF065436 Chlamydia
3	17.4	45.8	60	2	CTB1AOMP1	AF070239 Chlamydia
4	17.4	45.8	60	2	CTB2AOMP1	AF070275 Chlamydia
5	17.4	45.8	60	2	CTB2AOMP1	AF070255 Chlamydia
6	17.4	45.8	60	2	CTB2AOMP1	AF070271 Chlamydia
7	17.4	45.8	60	2	CTB2AOMP1	AF070251 Chlamydia
8	17.4	45.8	60	2	CTB2AOMP1	AF070251 Chlamydia
9	17.4	45.8	60	2	CTB2AOMP1	AF070259 Chlamydia
10	17.4	45.8	60	2	CTB2AOMP1	AF070263 Chlamydia
11	17.4	45.8	60	2	CTB2AOMP1	AF070267 Chlamydia
12	17.4	45.8	60	2	CTB2AOMP1	AF070279 Chlamydia
13	17.4	45.8	60	2	CTB2AOMP1	AF070283 Chlamydia
14	17.4	45.8	60	2	CTB2AOMP1	AF070291 Chlamydia
15	17.4	45.8	60	2	CTB2AOMP1	AF070295 Chlamydia
16	17.4	45.8	60	2	CTB2AOMP1	AF070299 Chlamydia
17	17.4	45.8	60	2	CTB2AOMP1	AF070303 Chlamydia
18	17.4	45.8	60	2	CTB2AOMP1	AF070235 Chlamydia
19	17.4	45.8	100	12	MA286084	286084 M. auratus D
20	16.6	43.7	45	5	AR012219	AR012219 Sequence
21	16.6	43.7	45	5	I21230	I21230 Sequence 4
22	16.6	43.7	45	5	I25875	I25875 Sequence 4
23	16.6	43.7	45	5	I26268	I26268 Sequence 4
24	16.4	43.2	60	5	A17027	A17027 oligonucleo
25	16.4	43.2	60	5	A17427	A17427 oligonucleo
26	16.4	43.2	60	5	A21213	A21213 Synthetic n
27	16.4	43.2	60	5	A23859	A23859 oligonucleo
28	16.4	43.2	60	5	AR014402	AR014402 Sequence
29	16.4	43.2	60	5	AR059804	AR059804 Sequence
30	16.4	43.2	60	5	AR060211	AR060211 Sequence
31	16.4	43.2	60	5	I11845	I11845 Sequence 11
32	16.4	43.2	81	3	PIGPROG	M25036 Hog prosta
33	16.4	43.2	92	10	HUMPHOS15	M59733 Human muscl
34	16.2	42.1	39	5	AR012248	AR012248 Sequence
35	16.2	42.1	39	5	I14946	I14946 Sequence 32
36	16.2	42.1	39	5	I73666	I73666 Sequence 32
37	16.2	42.1	73	4	CHKCSRC3	M61222 G.gallus c-
38	15.8	41.6	45	5	A47907	A47907 Sequence 4
39	15.8	41.6	60	2	AR028430	AR028430 Sequence
40	15.8	41.6	60	2	CTB12AOMP1	AF070247 Chlamydia
41	15.8	41.6	60	2	CTB12AOMP1	AF070243 Chlamydia
42	15.8	41.6	60	2	CTB12AOMP1	AF070231 Chlamydia
43	15.8	41.6	60	2	CTB12AOMP1	AF070227 Chlamydia
44	15.8	41.6	60	2	CTB12AOMP1	AF178225 Chlamydia
45	15.8	41.6	60	2	CTB12AOMP1	AF178229 Chlamydia

ALIGNMENTS

RESULT 1
LOCUS S77980 60 bp DNA BCT 27-SEP-1995
DEFINITION omp1-major outer membrane protein [Chlamydia trachomatis, genovar B variants B1/B3, Gambian isolate, conjunctival swab, Genomic, 60 nt].
ACCESSION S77980 GI:999268
VERSION S77980
KEYWORDS Chlamydia trachomatis Gambian isolate conjunctival swab genovar B
SOURCE variants B1/B3.
ORGANISM Chlamydia trachomatis
REFERENCE 1 (bases 1 to 60)
AUTHORS Hayes, L.J., Pecharatana, S., Bailey, R.L., Hampton, T.J., Pickett, M.A., Mabey, D.C., Watt, P.J. and Ward, M.E.
TITLE Extent and kinetics of genetic change in the omp1 gene of Chlamydia trachomatis in two villages with endemic trachoma
JOURNAL Infect. Dis. 172 (1), 268-272 (1995)
MEDLINE 95318541
REMARK Genbank staff at the National Library of Medicine created this entry [NCBI glibsg 166874] from the original journal article.
FEATURES
source Location/Qualifiers
1..60
/organism="Chlamydia trachomatis"
/db_xref="taxon:813"
1..60
/partial
/gene="omp1"
/note="major outer membrane protein, MOMP"
1..60
/partial
/gene="omp1"
/note="this sequence comes from Fig. 1A; MOMP"
/codon_start=1
/product="major outer membrane protein"
/protein_id="AAB34695.1"
/db_xref="GI:999269"
/translation="AKPTTGGNAAPSTLTARE"
BASE COUNT 19 a 20 c 11 g 10 t
ORIGIN

Query Match 45.8%; Score 17.4; DB 1; Length 60;
Best Local Similarity 68.6%; Pred. No. 3.6e+03;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 2 atgctgcagcttgacccaatttggcaccacacatag 36
|||||
Db 26 atgctgcagcttcacacactcttaccgacnagag 60

RESULT 2
AF065436 60 bp DNA BCT 23-MAY-1998
LOCUS Chlamydia trachomatis major outer membrane protein (omp-1) gene, partial cds.
DEFINITION AF065436
ACCESSION AF065436 GI:3152925
VERSION AF065436.1 GI:3152925
KEYWORDS Chlamydia trachomatis.
SOURCE Chlamydia trachomatis.
ORGANISM Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE 1 (bases 1 to 60)
AUTHORS Hsieh, Y.-H. and Bobo, L.D.
TITLE Diversity of major outer membrane protein (omp-1) gene of Chlamydia trachomatis in trachoma endemic villages
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 60)
AUTHORS Hsieh, Y.-H. and Bobo, L.D.
TITLE Direct Submission

JOURNAL Submitted (14-MAY-1998) Pediatric Infectious Diseases, Johns Hopkins Hospital, 600 N. Wolfe Street, Ballock 1138, Baltimore, MD 21287-4933, USA
FEATURES
source Location/Qualifiers
1..60
/organism="Chlamydia trachomatis"
/db_xref="taxon:813"
/note="genovar B"
1..>60
/gene="omp-1"
1..>60
/gene="omp-1"
/note="encodes variable domain"
/codon_start=1
/transl_table=1
/product="major outer membrane protein"
/protein_id="AAC17176.1"
/db_xref="GI:3152926"
/translation="AKPRATYGNAAAPSTLTARE"
BASE COUNT 18 a 20 c 12 g 10 t
ORIGIN

Query Match 45.8%; Score 17.4; DB 2; Length 60;
Best Local Similarity 68.6%; Pred. No. 3.6e+03;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 2 atgctgcagcttgacccaatttggcaccacacatag 36
|||||
Db 26 atgctgcagcttcacacactcttaccgacnagag 60

RESULT 3
CTB1AOMP1 60 bp DNA BCT 02-JUL-1998
LOCUS Chlamydia trachomatis strain B1a major outer membrane protein (omp-1) gene, partial cds.
DEFINITION AF070239
ACCESSION AF070239 GI:3283463
VERSION AF070239.1 GI:3283463
KEYWORDS 1 of 4
SEGMENT Chlamydia trachomatis.
SOURCE Chlamydia trachomatis.
ORGANISM Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE 1 (bases 1 to 60)
AUTHORS Hsieh, Y.-H. and Bobo, L.D.
TITLE Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 60)
AUTHORS Hsieh, Y.-H. and Bobo, L.D.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1998) Division of Pediatric Infectious Diseases, School of Medicine, Johns Hopkins University, Ballock 1138, 600 N. Wolfe Street, Baltimore, MD 21287-4933, USA
FEATURES
source Location/Qualifiers
1..60
/organism="Chlamydia trachomatis"
/strain="B1a"
/db_xref="taxon:813"
1..>60
/gene="omp-1"
/codon_start=1
/transl_table=1
/product="major outer membrane protein"
/protein_id="AAC25209.1"
/db_xref="GI:3283468"
/translation="ARPRATYGNAAAPSTLTARE"
BASE COUNT 17 a 20 c 13 g 10 t
ORIGIN

Query Match 45.8%; Score 17.4; DB 2; Length 60;

Best Local Similarity 68.6%; Pred. No. 3.6e+03;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 2 atgtcagcttgaccatttggcaccacacatag 36
|||||
Db 26 ATGCTGCAGCTCCATCCACTCTTACAGCAGAGAG 60

RESULT 4

CTB24AOMP1

LOCUS

DEFINITION Chlamydia trachomatis strain B24a major outer membrane protein

(omp-1) gene, partial cds.

ACCESSION AF070275.1 GI:3283544

VERSION

KEYWORDS

SEGMENT

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 60)

AUTHORS

TITLE

Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania

Unpublished

2 (bases 1 to 60)

AUTHORS

Hsieh, Y.-H. and Bobo, L.D.

TITLE

Submitted (02-JUN-1998) Division of Pediatric Infectious Diseases, School of Medicine, Johns Hopkins University, Blalock 1138, 600 N. Wolfe Street, Baltimore, MD 21287-4933, USA

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Unpublished
2 (bases 1 to 60)

Hsieh, Y.-H. and Bobo, L.D.

Direct Submission

Submitted (02-JUN-1998) Division of Pediatric Infectious Diseases, School of Medicine, Johns Hopkins University, Blalock 1138, 600 N. Wolfe Street, Baltimore, MD 21287-4933, USA

Location/Qualifiers

FEATURES

source

1. 60

/organism="Chlamydia trachomatis"

/strain="B24a"

/db_xref="taxon:813"

<1..>60

/gene="omp-1"

/codon_start=1

/transl_table=1

/product="major outer membrane protein"

/protein_id="AAC25225.1"

/db_xref="GI:3283504"

/translation="AKPTATTGNAAPSTLTARE"

BASE COUNT 18 a 20 c 12 g 10 t

ORIGIN

Query Match

45.8%; Score 17.4; DB 2; Length 60;

Best Local Similarity 68.6%; Pred. No. 3.6e+03;

Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 2 atgtcagcttgaccatttggcaccacacatag 36
|||||
Db 26 ATGCTGCAGCTCCATCCACTCTTACAGCAGAGAG 60

RESULT 6

CTB24BOMP1

LOCUS

DEFINITION Chlamydia trachomatis strain B24b major outer membrane protein

(omp-1) gene, partial cds.

ACCESSION AF070271.1 GI:3283535

VERSION

KEYWORDS

SEGMENT

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 60)

AUTHORS

TITLE

Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania

Unpublished

2 (bases 1 to 60)

AUTHORS

Hsieh, Y.-H. and Bobo, L.D.

TITLE

Submitted (02-JUN-1998) Division of Pediatric Infectious Diseases, School of Medicine, Johns Hopkins University, Blalock 1138, 600 N. Wolfe Street, Baltimore, MD 21287-4933, USA

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Query Match 45.8%; Score 17.4; DB 2; Length 60;
Best Local Similarity 68.6%; Pred. No. 3.6e+03;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 atgtctgacgttgaccactttggcaccacatag 36
|||||
Db 26 ATGCTGCAGCTCCATCCTTACGACGAGAG 60

RESULT 7
CTB2AOMP1 60 bp DNA BCT 02-JUL-1998
DEFINITION Chlamydia trachomatis strain B2a major outer membrane protein
(omp-1) gene, partial cds.
ACCESSION AF070251
VERSION AF070251.1 GI:3283490
KEYWORDS
SEGMENT
SOURCE
ORGANISM Chlamydia trachomatis.
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE
AUTHORS Hsieh, Y.-H. and Bobo, L.D.
TITLE Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania
JOURNAL Unpublished
REFERENCE
AUTHORS Hsieh, Y.-H. and Bobo, L.D.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1998) Division of Pediatric Infectious Diseases, School of Medicine, Johns Hopkins University, Blalock 1138, 600 N. Wolfe Street, Baltimore, MD 21287-4933, USA

FEATURES
source
1. 60
/organism="Chlamydia trachomatis"
/strain="B2a"
/db_xref="taxon:813"
<1. .>60
/gene="omp-1"
/codon_start=1
/transl_table=11
/product="major outer membrane protein"
/protein_id="AAC25221.1"
/db_xref="GI:3283495"
/translation="AKPYATGNAAPSTLTARE"
BASE COUNT 18 a 20 c 12 g 10 t
ORIGIN

Query Match 45.8%; Score 17.4; DB 2; Length 60;
Best Local Similarity 68.6%; Pred. No. 3.6e+03;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 atgtctgacgttgaccactttggcaccacatag 36
|||||
Db 26 ATGCTGCAGCTCCATCCTTACGACGAGAG 60

RESULT 8
CTB2BOMP1 60 bp DNA BCT 02-JUL-1998
DEFINITION Chlamydia trachomatis strain B2b major outer membrane protein
(omp-1) gene, partial cds.
ACCESSION AF070259
VERSION AF070259.1 GI:3283508
KEYWORDS
SEGMENT
SOURCE
ORGANISM Chlamydia trachomatis.
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE
AUTHORS Hsieh, Y.-H. and Bobo, L.D.
TITLE Diversity of major outer membrane protein (omp-1) of Chlamydia

JOURNAL Unpublished
REFERENCE
AUTHORS Hsieh, Y.-H. and Bobo, L.D.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1998) Division of Pediatric Infectious Diseases, School of Medicine, Johns Hopkins University, Blalock 1138, 600 N. Wolfe Street, Baltimore, MD 21287-4933, USA

FEATURES
source
1. 60
/organism="Chlamydia trachomatis"
/strain="B2b"
/db_xref="taxon:813"
<1. .>60
/gene="omp-1"
/codon_start=1
/transl_table=11
/product="major outer membrane protein"
/protein_id="AAC25223.1"
/db_xref="GI:3283513"
/translation="AKPYATGNAAPSTLTARE"
BASE COUNT 18 a 20 c 12 g 10 t
ORIGIN

Query Match 45.8%; Score 17.4; DB 2; Length 60;
Best Local Similarity 68.6%; Pred. No. 3.6e+03;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 atgtctgacgttgaccactttggcaccacatag 36
|||||
Db 26 ATGCTGCAGCTCCATCCTTACGACGAGAG 60

RESULT 9
CTB2COMP1 60 bp DNA BCT 02-JUL-1998
DEFINITION Chlamydia trachomatis strain B2c major outer membrane protein
(omp-1) gene, partial cds.
ACCESSION AF070263
VERSION AF070263.1 GI:3283517
KEYWORDS
SEGMENT
SOURCE
ORGANISM Chlamydia trachomatis.
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE
AUTHORS Hsieh, Y.-H. and Bobo, L.D.
TITLE Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania
JOURNAL Unpublished
REFERENCE
AUTHORS Hsieh, Y.-H. and Bobo, L.D.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1998) Division of Pediatric Infectious Diseases, School of Medicine, Johns Hopkins University, Blalock 1138, 600 N. Wolfe Street, Baltimore, MD 21287-4933, USA

FEATURES
source
1. 60
/organism="Chlamydia trachomatis"
/strain="B2c"
/db_xref="taxon:813"
<1. .>60
/gene="omp-1"
/codon_start=1
/transl_table=11
/product="major outer membrane protein"
/protein_id="AAC25233.1"
/db_xref="GI:3283523"
/translation="AKPYATGNAAPSTLTARE"
BASE COUNT 18 a 20 c 12 g 10 t
ORIGIN

Query Match 45.8%; Score 17.4; DB 2; Length 60;
Best Local Similarity 68.6%; Pred. No. 3.6e+03;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 atgtctgacgttgaccatttggcaccacacatag 36
|||||
Db 26 ATGCTGCAGCTCCATCCACTTACAGCAGAGAG 60

RESULT 10
CTB2DOMP1 60 bp DNA BCT 02-JUL-1998
DEFINITION Chlamydia trachomatis strain B2d major outer membrane protein
(omp-1) gene, partial cds.
ACCESSION AF070267
VERSION AF070267.1 GI:3283526
KEYWORDS
SEGMENT
SOURCE 1 of 4
ORGANISM Chlamydia trachomatis.
REFERENCE Chlamydia trachomatis
AUTHORS Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
TITLE 1 (bases 1 to 60)
Hsieh,Y.-H. and Bobo,L.D.
JOURNAL Diversity of major outer membrane protein (omp-1) of Chlamydia
REFERENCES trachomatis in trachoma endemic villages, Kongwa, Tanzania
AUTHORS Unpublished
TITLE 2 (bases 1 to 60)
Hsieh,Y.-H. and Bobo,L.D.
JOURNAL Submitted (02-JUN-1998) Division of Pediatric Infectious Diseases,
REFERENCES School of Medicine, Johns Hopkins University, Ballock 1138, 600 N.
Wolfe Street, Baltimore, MD 21287-4933, USA
LOCATION/Qualifiers
source 1. 60
/organism="Chlamydia trachomatis"
/strain="B2d"
/db_xref="taxon:813"
<1. >60
/gene="omp-1"
/codon_start=1
/transl_table=11
/product="major outer membrane protein"
/protein_id="AAC25237.1"
/db_xref="GI:3283531"
/translation="AKPTATGTGNAAPSTLTARE"
BASE COUNT 18 a 20 c 12 g 10 t
ORIGIN

Query Match 45.8%; Score 17.4; DB 2; Length 60;
Best Local Similarity 68.6%; Pred. No. 3.6e+03;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 atgtctgacgttgaccatttggcaccacacatag 36
|||||
Db 26 ATGCTGCAGCTCCATCCACTTACAGCAGAGAG 60

RESULT 11
CTB2EOMP1 60 bp DNA BCT 02-JUL-1998
DEFINITION Chlamydia trachomatis strain B2e major outer membrane protein
(omp-1) gene, partial cds.
ACCESSION AF070279
VERSION AF070279.1 GI:3283553
KEYWORDS
SEGMENT
SOURCE 1 of 4
ORGANISM Chlamydia trachomatis.
REFERENCE Chlamydia trachomatis.
AUTHORS Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
TITLE 1 (bases 1 to 60)
Hsieh,Y.-H. and Bobo,L.D.

TITLE Diversity of major outer membrane protein (omp-1) of Chlamydia
JOURNAL trachomatis in trachoma endemic villages, Kongwa, Tanzania
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 60)
TITLE Hsieh,Y.-H. and Bobo,L.D.
JOURNAL Direct Submission
SUBMITTED (02-JUN-1998) Division of Pediatric Infectious Diseases,
SCHOOL OF MEDICINE, Johns Hopkins University, Ballock 1138, 600 N.
Wolfe Street, Baltimore, MD 21287-4933, USA
LOCATION/Qualifiers
source 1. 60
/organism="Chlamydia trachomatis"
/strain="B2e"
/db_xref="taxon:813"
<1. >60
/gene="omp-1"
/codon_start=1
/transl_table=11
/product="major outer membrane protein"
/protein_id="AAC25249.1"
/db_xref="GI:3283558"
/translation="AKPTATGTGNAAPSTLTARE"
BASE COUNT 18 a 20 c 12 g 10 t
ORIGIN

Query Match 45.8%; Score 17.4; DB 2; Length 60;
Best Local Similarity 68.6%; Pred. No. 3.6e+03;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 atgtctgacgttgaccatttggcaccacacatag 36
|||||
Db 26 ATGCTGCAGCTCCATCCACTTACAGCAGAGAG 60

RESULT 12
CTB2FOMP1 60 bp DNA BCT 02-JUL-1998
DEFINITION Chlamydia trachomatis strain B2f major outer membrane protein
(omp-1) gene, partial cds.
ACCESSION AF070283
VERSION AF070283.1 GI:3283562
KEYWORDS
SEGMENT
SOURCE 1 of 4
ORGANISM Chlamydia trachomatis.
REFERENCE Chlamydia trachomatis.
AUTHORS Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
TITLE 1 (bases 1 to 60)
Hsieh,Y.-H. and Bobo,L.D.
JOURNAL Diversity of major outer membrane protein (omp-1) of Chlamydia
REFERENCES trachomatis in trachoma endemic villages, Kongwa, Tanzania
AUTHORS Unpublished
TITLE 2 (bases 1 to 60)
Hsieh,Y.-H. and Bobo,L.D.
JOURNAL Direct Submission
SUBMITTED (02-JUN-1998) Division of Pediatric Infectious Diseases,
SCHOOL OF MEDICINE, Johns Hopkins University, Ballock 1138, 600 N.
Wolfe Street, Baltimore, MD 21287-4933, USA
LOCATION/Qualifiers
source 1. 60
/organism="Chlamydia trachomatis"
/strain="B2f"
/db_xref="taxon:813"
<1. >60
/gene="omp-1"
/codon_start=1
/transl_table=11
/product="major outer membrane protein"
/protein_id="AAC25253.1"
/db_xref="GI:3283567"
/translation="AKPTATGTGNAAPSTLTARE"
BASE COUNT 18 a 20 c 12 g 10 t
ORIGIN

Tue Jun 6 08:33:49 2000

us-09-164-714-18.rge

Page 7

ORIGIN

Query Match	45.8%;	Score 17.4;	DB 2;	Length 60;
Best Local Similarity	68.6%;	Pred. No. 3.6e+03;		
Matches 24;	Conservative	0;	Mismatches 11;	IndeIs 0;
				Gaps 0;

```

Qy      2 atgtcgagcttgaccaatttggcaccacaatag 36
          |||||  |  |  |  |  |  |  |  |  |
Db      26 ATGTCGAGCTTCATCCACTCTTACAGCAGAGAG 60

```

Search completed: June 4, 2000, 16:05:44
Job time: 27896 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:20 ; Search time 322.35 Seconds
(without alignments)
29.494 Million cell updates/sec

Title: US-09-164-714-18

Perfect score: 38
Sequence: 1 catgctgcagcttgcacacatttggcacaacataaggg 38

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	19.6	51.6	96	1	Q40167	Sequence encoding
C 2	19.6	51.6	96	1	Q46018	CDNA derived from
C 3	19.6	51.6	96	1	Q47688	Encodes antiviral
C 4	19.6	51.6	96	1	Q56287	Epstein-Barr virus
C 5	19.6	51.6	96	1	Q75265	5'-terminal of ins
6	17.4	45.8	60	1	N92143	Chlamydia trachoma
7	17.4	45.8	60	1	N92147	Chlamydia trachoma
8	17.4	45.8	60	1	N97061	Sequence of C. tra
9	17.4	45.8	60	1	N97065	Sequence of C. tra
10	17.4	45.8	63	1	T28244	Feline herpesvirus
C 11	16.8	44.2	33	1	V45319	Heavy chain forwar
C 12	16.6	43.7	45	1	Q85044	3-Hydroxysteroid-o
C 13	16.6	43.7	45	1	T62671	Streptomyces sp. s
C 14	16.6	43.7	45	1	V11816	Human granulocyte
C 15	16.4	43.2	41	1	X19304	Feline herpesvirus
16	16.4	43.2	90	1	T28251	Human granulocyte
17	16.2	42.6	41	1	X19305	Human granulocyte
18	15.8	41.6	24	1	X22061	PCR primer for C.
C 19	15.8	41.6	31	1	X38735	Human genomic DNA
C 20	15.8	41.6	45	1	T10346	CAEV env gene TM f
C 21	15.8	41.6	60	1	N92131	Chlamydia trachoma
22	15.8	41.6	60	1	N92139	Sequence of Chlamy
23	15.8	41.6	60	1	N97049	Sequence of C. tra
24	15.8	41.6	60	1	N97057	Sequence of C. tra
25	15.8	41.6	60	1	X18016	C. trachomatis ser
26	15.8	41.6	60	1	X18008	C. trachomatis ser
27	15.8	41.6	61	1	N92127	Sequence of Chlamy
28	15.8	41.6	61	1	N97045	Sequence of C. tra
C 29	15.6	41.1	50	1	Q34114	Downstream sequenc
C 30	15.6	41.1	50	1	V76355	Staphylococcus aur
C 31	15.6	41.1	70	1	X16757	Alpha-helix peptid
C 32	15.6	41.1	71	1	V76550	Staphylococcus aur
C 33	15.4	40.5	50	1	Q12961	Template linker pr
34	15.4	40.5	50	1	T74411	HIV template linke

35	15.4	40.5	63	1	T28230	Feline herpesvirus
36	15.2	40.0	29	1	V43049	5' PCR primer Deg-
C 37	15.2	40.0	49	1	V13138	Oligonucleotide pr
C 38	15.2	40.0	51	1	T05046	LK26 heavy chain v
C 39	15.2	40.0	85	1	Q27291	TGF-alpha construc
C 40	15.2	40.0	85	1	Q44603	Plasmid pBS114 con
C 41	15.2	40.0	86	1	Q27293	TGF-alpha construc
42	15.2	40.0	86	1	Q44605	Plasmid pBS114 con
C 43	15	39.5	34	1	Q71835	Primer FIB2. fibro
C 44	15	39.5	34	1	T18304	Primer for amplifi
C 45	15	39.5	85	1	X04417	Fragment of human

ALIGNMENTS

```
RESULT 1
ID Q40167/c
AC Q40167;
DE 09-AUG-1993 (first entry)
KW HIV-1; HIV-2; HHV-6; therapy; antiviral; AIDS; ss.
OS Sus scrofa domestica.
PN EP-537722-A.
PD 21-APR-1993.
PF 14-OCT-1992; 117569.
PR 15-OCT-1991; US-780084.
PR 04-DEC-1991; US-803844.
PA (KREM-) KREMERS-URBAN CO.
PI Gawish AAS, Hermann WJ, Lawyer CH, Steinbach T,
PI Tanaka SK, Magle SS;
DR WPI; 93-127605/16.
PT Polypeptide obtained from a porcine liver extract - useful for
PT treating HIV-1, HIV-2, HIV-6 and AIDS infections
PS Claim 6; Page 25; 42pp; English.
CC The inventors claim a novel polypeptide with antiviral activity and
CC a mol. wt. of 5-49 KD obtd. from an acetone-insoluble liver extract.
CC The polypeptide has 5' and 3' amino acid sequences selected from
CC those given in the specification encoded by DNA sequences also given
CC in the specification. Synthetic peptides are also claimed. The
CC polypeptides are used in compns. for the treatment of viral
CC infections, partic. HIV-1, HIV-2 and human herpes virus (HHV-6)
CC infections. The polypeptides are capable of inhibiting the
CC replication and pathogenesis of the viruses.
SQ Sequence 96 BP; 22 A; 21 C; 30 T;

Query Match 51.6%; Score 19.6; DB 1; Length 96;
Best Local Similarity 73.5%; Pred. No. 21;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Oy 2 atgctgcagcttgcacacatttggcacaacata 35
Db 79 AAGCTGCAGCTTGACGTATTTATTCACCTTGATA 46

RESULT 2
ID Q46018/c
AC Q46018;
DE 24-JAN-1994 (first entry)
KW Alzheimer's disease; dementia; medicament; chronic fatigue syndrome;
KW peptide; Kutapressin; ss.
OS Sus scrofa.
PN EP-554887-A.
PD 11-AUG-1993.
PF 05-FEB-1993; 101819.
PF 05-FEB-1992; US-835029.
PA (KREM-) KREMERS-URBAN CO.
PI Gawish AAS, Hermann WJ, Lawyer CH, Steinbach T;
PI Magle SS;
```

DR WPI: 93-251162/32.
 DR P-PSDB: R39512.
 PT Medical use of liver extract fraction and/or peptide(s) - for
 PS prodn. of medicaments for treating pre-senile or senile dementia
 CC Claim 1; Page 12; 199P; English.
 CC A heat stable acetone-insoluble, water-soluble mammalian liver
 CC extract is used to treat patients with alzheimers disease since
 CC it is effective in the treatment of mammals infected with non-
 CC dermatologic viruses, in particular, chronic fatigue syndrome, the
 CC symptoms of which are very similar to a mild form of nonprogressive
 CC acquired Alzheimer like disease. Therapeutically effective amounts
 CC of a specific peptide or peptide fragments (R39512-R39516) can also
 CC be used to treat Alzheimers disease. The peptide is purified
 CC from the liver extract and is administered preferably by
 CC intramuscular injection. The CDNA sequences encoding this
 CC peptide/fragments was derived from rat liver RNA (Q46018, Q46020-21).
 SQ Sequence 96 BP; 22 A; 20 C; 24 G; 30 T;

Query Match 51.6%; Score 19.6; DB 1; Length 96;
 Best Local Similarity 73.5%; Pred. No. 21;
 Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 2 atcgtcagcttgaccatttggcaccacata 35
 Db 79 AAGCTGACCTTGACGTATATTGCACTTGATA 46

RESULT 3
 O47688/c
 ID O47688 standard; DNA: 96 BP.
 AC O47688;
 DE 08-FEB-1994 (first entry)
 DE Encodes antiviral peptide N-terminal peptide from pig liver.
 KW anti-viral polypeptide; HIV; human immunodeficiency virus; HHV-6;
 KW human herpes virus-6; AIDS; acquired immune deficiency syndrome;
 KW AZT synergist; mammalian liver; porcine; insoluble extract;
 KW Kutapressin.
 OS Sus scrofa domestica.
 FH Key Location/Qualifiers
 FT misc_feature 1..96
 FT /tag= a
 FT /note= "partial open reading frame"
 FT misc_difference 76..78
 FT /tag= b
 FT /note= "GCT codon is shown as encoding Pro"

CA2084331-A.
 PD 05-JUN-1993
 PR 02-DEC-1992; 084331.
 PR 04-DEC-1991; US-803844.
 PA 14-OCT-1992; US-962570.
 PA (MAGL/) MAGLE S S.
 PI Gawish AAS, Hermann WJ, Lawyer CH, Steinbach T;
 PI Tanaka SK, Magle SS;
 DR WPI: 93-265125/34.
 DR P-PSDB: R40775.
 PT New polypeptide(s) from acetone-insol. liver extract - have
 PT antiviral activity, partic. for treatment of HIV and human
 PT herpes virus-6 infections
 PS Claim 4; Page 51; 72pp; English.
 CC Polypeptides with antiviral activity and mol.wt. 5000-40000
 CC were isolated from the acetone-insoluble fraction of commercially
 CC available mammalian liver extract "KUTAPRESSIN" (Kremers-Urban Co.).
 CC Preferred polypeptides have the amino acid sequence R40775 at the
 CC N-terminal and either R40776 or R40777 at the C-terminal. Genomic
 CC DNA sequences encoding such polypeptides were isolated using PCR
 CC primers based on partial replication fragments. The polypeptides were
 CC found to inhibit viral replication and to be effective against
 CC HHV-6, HIV-1 and HIV-2. They can act synergistically with AZT to
 CC suppress HIV infection and AIDS.
 SQ Sequence 96 BP; 22 A; 20 C; 24 G; 30 T;

Query Match 51.6%; Score 19.6; DB 1; Length 96;
 Best Local Similarity 73.5%; Pred. No. 21;
 Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 2 atcgtcagcttgaccatttggcaccacata 35
 Db 79 AAGCTGACCTTGACGTATATTGCACTTGATA 46

RESULT 4
 O56287/c
 ID O56287 standard; DNA: 96 BP.
 AC O56287;
 DE 02-AUG-1994 (first entry)
 DE Epstein-Barr virus infection therapeutic peptide gene.
 KW Epstein-Barr virus; EBV; chronic infectious mononucleosis;
 KW liver; KU10,001.
 OS Sus scrofa.
 FH Key Location/Qualifiers
 FT cds 1..96
 FT /tag= a
 FT /note= "5' region of peptide gene"

MO9404170-A.
 PD 03-MAR-1994.
 PR 23-AUG-1993; U07877.
 PR 24-AUG-1992; US-934554.
 PA (KREM-) KREMERS-URBAN CO.
 PI Gawish AAS, Hermann WJ, Lawyer CH, Steinbach T;
 PI Magle SS;
 DR WPI: 94-082824/10.
 DR P-PSDB: R49251.
 PT Method of treating Epstein Barr virus infectious, e.g. chronic
 PT infectious mononucleosis - by admin. porcine liver extracts
 PT KU10,001 or peptides derived from this
 PS Disclosure; Page 22; 39pp; English.
 CC DNA encoding a pig liver extract (KU10,001) peptide active against
 CC EBV infection was sequenced. The 5' sequence (O56287) encoded the
 CC peptide given in R49251. The 3' region had the sequence given in
 CC O56288. The active peptide C-terminus either had sequence R49252
 CC (encoded by O56289, a part of O56288) or R49253. The combination
 CC of R49251 with R49252 gave peptide R49254, and with R49253 gave
 CC R49255.
 SQ Sequence 96 BP; 22 A; 20 C; 24 G; 30 T;

Query Match 51.6%; Score 19.6; DB 1; Length 96;
 Best Local Similarity 73.5%; Pred. No. 21;
 Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 2 atcgtcagcttgaccatttggcaccacata 35
 Db 79 AAGCTGACCTTGACGTATATTGCACTTGATA 46

RESULT 5
 O75265/c
 ID O75265 standard; CDNA: 96 BP.
 AC O75265;
 DE 19-JAN-1995 (first entry)
 DE 5'-terminal of insert coding for porcine antiviral polypeptide.
 KW Hepatitis B virus; infection; treatment; porcine polypeptide;
 KW kutapressin (RTM); KU 10,001; antiviral activity; ds.
 OS Sus scrofa.
 PN US5316775-A.
 PD 31-MAY-1994.
 PR 04-AUG-1988; 228364.
 PR 19-APR-1993; US-049511.
 PA (KREM-) KREMERS-URBAN CO.
 PI Gawish AAS, Hermann WJ, Lawyer CH, Steinbach T;
 PI Magle SS;
 DR WPI: 94-176217/21.
 DR P-PSDB: R66493.
 PT Treating hepatitis B infection by admin. of mammalian liver

PT extract - which is heat stable and insoluble in acetone but
PT soluble in water.
PS Example 4; Columns 13-14; 25pp; English.
CC Polypeptides purified from mammalian liver extract which showed
CC cell protection activity when tested for anti-viral activity were
CC partially sequenced. A degenerate oligonucleotide primer (075269)
CC was designed based on the similarities between the different peptide
CC sequences (R66498-R66501) from active polypeptides KU 10,214 and KU
CC 10,215. The degenerate primer was used for PCR on cDNA prepared from
CC pig liver poly (A)+ RNA. Amplified DNA fragments were analysed. One
CC insert of 500bp was found in six out of four randomly chosen isolates.
CC DNA sequencing determined 96bp at the 5'-end and 110bp at the 3'-end
CC of the insert (see 075265 and 075266, respectively). The 3'-end had
CC termination codons in all three reading frames giving 3 alternative
CC C-terminal sequences (i.e. Leu, R66494 or R66495). The active
CC polypeptides are thus characterised by sequences R66496 or R66497.
SQ Sequence 96 BP; 22 A; 20 C; 24 G; 30 T;

Query Match 51.6%; Score 19.6; DB 1; Length 96;
Best Local Similarity 73.5%; Pred. No. 21;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Oy 2 atgtcgcagcttgacatttgccacacata 35
| ||||| ||||| ||||| ||||| |||||
Db 79 AAGCTGCAGCTTGACGATTTATTCACCTTGCAATA 46

RESULT 6
ID N92143
AC N92143;
DT 12-APR-1990 (first entry)
DE Chlamydia trachomatis serovars L1 major outer membrane protein
DE variable domain L1-VDI.
KW Chlamydia trachomatis serovars L1; variable domain L1-VDI;
KW major outer membrane protein; monoclonal antibody; vaccine.
OS Chlamydia trachomatis serovars L1.
FH Key Location/Qualifiers
FT cds 1..90
FT /*tag= a
PN US7324664-A.
PD 29-AUG-1989.
PF 17-MAR-1989; 324664.
PR 17-MAR-1989; US-324664.
PA (USSH) US Dept. of Health and Human Services.
PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
PI WPI: 89-339697/46.
DR P-PSDB; P93319.
PT Chlamydia trachomatis genes - used for determ. of nucleotide and amino
PT sequences of the variable domains of the major outer membrane proteins.
PS Disclosure; Fig. 9; 49pp; English.
CC The major neutralising and serotyping antigenic determinants have been
CC located in the variable domain. The sequence was determined using
CC oligonucleotide primer extension sequencing. The variable domain is
CC recognised by MOMP specific monoclonal antibodies. The domain will be
CC useful in the construction of synthetic peptides, subunits or recombinant
CC chlamydial vaccines. They will also allow prodn. of reagents and methods
CC for development of new diagnostic tests for C. trachomatis infections
CC and serological tests for serotyping. The domain is bases 256-315 of the
CC MOMP gene.
SQ Sequence 60 BP; 18 A; 19 C; 13 G; 10 T;

Query Match 45.8%; Score 17.4; DB 1; Length 60;
Best Local Similarity 68.6%; Pred. No. 1.5e+02;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Oy 2 atgtcgcagcttgacatttgccacacata 36
| ||||| ||||| ||||| ||||| |||||
Db 26 ATGCTGCAGCTTGACGATTTATTCACGCAAGAGAG 60

RESULT 7
ID N92147
AC N92147;
DT 17-APR-1990 (first entry)
DE Chlamydia trachomatis serovars L2 major outer membrane protein
DE variable domain L2-VDI.
KW Chlamydia trachomatis serovars L2; variable domain L2-VDI;
KW major outer membrane protein; monoclonal antibody; vaccine.
OS Chlamydia trachomatis serovars L2.
FH Key Location/Qualifiers
FT cds 1..60
FT /*tag= a
PN US7324664-A.
PD 29-AUG-1989.
PF 17-MAR-1989; 324664.
PR 17-MAR-1989; US-324664.
PA (USSH) US Dept. of Health and Human Services.
PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
PI WPI: 89-339697/46.
DR P-PSDB; P91445.
PT Chlamydia trachomatis genes - used for determ. of nucleotide and amino
PT sequences of the variable domains of the major outer membrane proteins.
PS Disclosure; Fig. 10; 49pp; English.
CC The major neutralising and serotyping antigenic determinants have been
CC located in the variable domain. The sequence was determined using
CC oligonucleotide primer extension sequencing. The variable domain is
CC recognised by MOMP specific monoclonal antibodies. The domain will be
CC useful in the construction of synthetic peptides, subunits or recombinant
CC chlamydial vaccines. They will also allow prodn. of reagents and methods
CC for development of new diagnostic tests for C. trachomatis infections
CC and serological tests for serotyping. The domain is bases 256-315 of the
CC MOMP gene.
SQ Sequence 60 BP; 18 A; 19 C; 13 G; 10 T;

Query Match 45.8%; Score 17.4; DB 1; Length 60;
Best Local Similarity 68.6%; Pred. No. 1.5e+02;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Oy 2 atgtcgcagcttgacatttgccacacata 36
| ||||| ||||| ||||| ||||| |||||
Db 26 ATGCTGCAGCTTGACGATTTATTCACGCAAGAGAG 60

RESULT 8
ID N97061
AC N97061;
DT 06-MAR-1992 (first entry)
DE Sequence of C. trachomatis serovar L1 major outer membrane protein (MOMP)
DE variable domain (VD) gene L1-VDI base pairs 256-315
KW Chlamydia trachomatis; antigen; monoclonal antibody; vaccine;
KW diagnosis; serotyping; non-immunologic assay; ss.
OS Chlamydia trachomatis.
FH Key Location/Qualifiers
FT cds 1..60
FT /*tag= a
PN US7324664-A.
PD 29-AUG-1989.
PF 17-MAR-1989; 324664.
PR 17-MAR-1989; US-324664.
PA (USSH) US DEPT HEALTH & HUMAN.
PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
PI WPI: 89-339697/46.
DR P-PSDB; P98426.
PT Chlamydia trachomatis genes - used for determ. of nucleotide and
PT amino sequences of the variable domains of the major outer
PT membrane proteins
PS Disclosure; Fig. 9; 49pp; English.
CC The inventors sequenced the 4 MOMP VDs of ten C. trachomatis
CC serovars and the amino acid sequences were deduced. The MOMP VDs
CC with the greatest total hydrophilicity and charge values were found

CC to be the location of antigenic determinants recognised by MOMP
 CC specific monoclonal antibodies. The nucleotide, amino acid
 CC sequences and hydrophilicity/charge value analyses will assist in
 CC the selection of appropriate MOMP antigenic determinants to be used
 CC in the construction of synthetic peptides, subunits or recombinant
 CC chlamydial vaccines. This will allow the prodn. or reagents and
 CC methodologies applicable in the development of new diagnostic tests
 CC for serotyping.
 SQ Sequence 60 BP; 18 A; 19 C; 13 G; 10 T;

Query Match 45.8%; Score 17.4; DB 1; Length 60;
 Best Local Similarity 68.6%; Pred. No. 1.5e+02;
 Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 2 atgcctgacgttgcacatttggcaccaacatag 36
 |||||
 DB 26 ATGCTGACAGCTCCATCTGACAGCAAGAGAG 60

RESULT 9

ID N97065 standard; DNA: 60 BP.
 AC N97065;
 DT 06-MAR-1992 (first entry)
 DE Sequence of C. trachomatis serovar L2 major outer membrane protein (MOMP)
 DE variable domain (VD) gene L2-VDI base pairs 256-315
 KW Chlamydia trachomatis; antigen; monoclonal antibody; vaccine;
 KW diagnosis; serotyping; non-immunologic assay; ss.
 OS Chlamydia trachomatis.
 FH Key
 FT Location/Qualifiers
 FT cds 1..60
 FT US7324664-A.
 PN 29-AUG-1989.
 PF 17-MAR-1989; 324664.
 PR 17-MAR-1989; US-324664.
 PA (USSH) US DEPT HEALTH & HUMAN.
 PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
 DR WPI: 89-339697/46.
 DR P-PSDB: p98430.

PT Chlamydia trachomatis genes - used for determ. of nucleotide and
 PT amino sequences of the variable domains of the major outer
 PT membrane proteins
 PS Disclosure; Fig 10; 49pp; English.
 CC The inventors sequenced the 4 MOMP VDS of ten C. trachomatis
 CC serovars and the amino acid sequences were deduced. The MOMP VDS
 CC with the greatest total hydrophilicity and charge values were found
 CC to be the location of antigenic determinants recognised by MOMP
 CC specific monoclonal antibodies. The nucleotide, amino acid
 CC sequences and hydrophilicity/charge value analyses will assist in
 CC the selection of appropriate MOMP antigenic determinants to be used
 CC in the construction of synthetic peptides, subunits or recombinant
 CC chlamydial vaccines. This will allow the prodn. or reagents and
 CC methodologies applicable in the development of new diagnostic tests
 CC for serotyping.
 SQ Sequence 60 BP; 18 A; 19 C; 13 G; 10 T;

Query Match 45.8%; Score 17.4; DB 1; Length 60;
 Best Local Similarity 68.6%; Pred. No. 1.5e+02;
 Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 2 atgcctgacgttgcacatttggcaccaacatag 36
 |||||
 DB 26 ATGCTGACAGCTCCATCTGACAGCAAGAGAG 60

RESULT 10

ID T28244 standard; DNA: 63 BP.
 AC T28244;
 DT 18-DEC-1996 (first entry)

DE Feline herpesvirus homology vector 669-42.04 junction H fragment.
 KW Feline herpesvirus; FHV; deletion; insertion; glycoprotein E;
 KW gp E; unique short region; attenuated virus; vaccine; protection;
 KW cat; vector; delivery; antigen; homology vector; junction H; ds.
 OS Synthetic.

FH Key
 FT Location/Qualifiers
 FT misc_feature 1..13
 FT /tag= a
 FT /label= pseudorabies_virus (Bam HI #10)
 FT misc_feature 14..42
 FT /tag= b
 FT /label= synthetic_linker
 FT misc_feature 43..63
 FT /tag= c
 FT /label= feline_herpesvirus (Eco RI E)

PN W09613575-A1.
 PD 09-MAY-1996.
 PF 26-OCT-1995; U13975.
 PR 26-OCT-1994; US-329883.
 PA (SYTR) SYNTRO CORP.
 PI Cochran MD, McDonnell MW;
 DR WPI: 96-239489/24.

PT Recombinant feline herpes virus attenuated by alteration of the ge
 PT gene - and related homology vectors, useful in vaccines, also as
 PT vectors for delivering antigens or therapeutic agents to mammals and
 PT birds
 PS Claim 49: Fig 12E; 162pp; English.
 CC The present sequence is a junction H fragment from the claimed
 CC feline herpesvirus (FHV) homology vector 669-42.04, which was
 CC constructed for the purpose of deleting a portion of the FHV
 CC glycoprotein E (gp E) gene coding region, and inserting a foreign
 CC DNA. It incorporates an E. coli beta-galactosidase gene, and a
 CC feline immunodeficiency virus (FIV) protease (gag) gene flanked by
 CC FHV DNA. The foreign genes were inserted into the NotI site in the
 CC FHV homology vector 644-09.04.
 CC Altering the gp E gene of FHV, so that it no longer produces
 CC functional gp E, attenuates the virus. The attenuated virus is
 CC useful in vaccines to protect cats against FHV (and opt. against
 CC other diseases if appropriate DNA is inserted), and as a vector
 CC for delivering vaccinating antigens or therapeutic agents (e.g.
 CC antisense mols., ribozymes, interferon inducers, hormones,
 CC lymphokines, etc.) to other mammals (including humans) or birds.
 SQ Sequence 63 BP; 11 A; 22 C; 20 G; 10 T;

Query Match 45.8%; Score 17.4; DB 1; Length 63;
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;
 Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 gctgcagcttgcacatttggcacca 30
 |||||
 DB 36 GCTGACAGCTCCATCTGACAGCA 62

RESULT 11

ID V45319/C standard; DNA: 33 BP.
 AC V45319;
 DT 27-OCT-1998 (first entry)

DE Heavy chain forward primer.
 KW J chain; targeting molecule; epithelial; beta-sheet; asthma;
 KW cancer; inflammatory disorder; autoimmune disorder; celiac disease;
 KW colitis; pneumonia; cystic fibrosis; ss; PCR; primer; amplification.
 OS Synthetic.
 PN W09830592-A1.
 PD 16-JUL-1998.
 PF 09-JAN-1998; U00542.
 PR 10-JAN-1997; US-782481.
 PA (EPIC-) EPICYTE PHARM INC.
 PI Fitchen JH, Hein MB, Hlatk AC;
 DR WPI: 98-399067/34.
 PT New epithelial tissue targeting agent - used to deliver
 PT biologically active compounds to an epithelial surface for

```
PT Internalisation
PS Example 1; Page 61; 142pp; English.
CC The primers W4513-V45322 are used in the method of the invention
CC to synthesize targeting molecules (TM). The TMs are used to target
CC biological agents to epithelial surfaces at which they can be
CC internalised. The TMs comprise a polypeptide that: (a) forms a closed
CC covalent loop; (b) contains at least 3, preferably 4, peptide domains
CC having beta-sheet character separated by domains lacking beta-sheet
CC character; and (c) is not full length dimeric IgA. The TMs are useful to
CC prevent and/or treat diseases associated with epithelial surfaces, e.g.
CC asthma, cancer, (myco)bacterial, viral or fungal infection, inflammatory
CC disorders, autoimmune disorders, celiac disease, colitis, pneumonia
CC and cystic fibrosis.
SQ Sequence 33 BP; 7 A; 5 C; 9 G; 7 T;

Query Match 44.2%; Score 16.6; DB 1; Length 33;
Best Local Similarity 64.3%; Pred. No. 2.3e+02;
Matches 18; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 tgcgtcagcttgaccatttgaccaca 30
   1:|||||:|:|:|:|:|:|:|:|:|
DB 28 TSCTGAGYTKSACCTGATCTGTACCA 1

RESULT 12
OS 085044
ID 085044 standard; DNA; 45 BP.
AC 085044;
DT 02-SEP-1995 (first entry)
DE 3-Hydroxysteroid-oxidase peptide N2-derived probe.
KM 3-Hydroxysteroid-oxidase: N2; DNA probe; insecticide;
KW transgenic plant; insect control; ss.
OS Synthetic.
PN W09501098-A.
PD 12-JAN-1995.
PF 24-JUN-1994; U07252.
PR 28-JUN-1993; US-083948.
PA (MONS ) MONSANTO CO.
PI Corbin DR, Greenplate JT, Jennings MG, Purcell JP,
PI Sammons RD;
DR WPI: 95-060750/08.
PT Controlling lepidopteran insect infestation of plants - by
PT providing a 3-hydroxy;steroid oxidase for ingestion by the insect.
PS Disclosure; Page 13; 39pp; English.
CC Hybridization probe N2 is used with probes C1 (085045) and C2
CC (085046) on Southern blots of Streptomyces sp. A19249 genomic DNA
CC for isolation and cloning of the 3-hydroxysteroid-oxidase (3-HSO)
CC gene. This sequence may be used to construct a cotton
CC (Gossypium hirsutum) or maize (Zea mays) transgenic plant
CC expressing 3-HSO. Alternatively, the sequence may be
CC expressed in plant-colonizing microorganisms or in Escherichia
CC coli, and applied directly to the plant, for controlling
CC lepidopteran insect infestations.
SQ Sequence 45 BP; 8 A; 14 C; 15 G; 8 T;

Query Match 43.7%; Score 16.6; DB 1; Length 45;
Best Local Similarity 71.0%; Pred. No. 3e+02;
Matches 22; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 atgctgacgttgaccatttgaccaca 32
   1:|||||:|:|:|:|:|:|:|:|:|
DB 13 ATGCTGAGATGGCCAGCTGTGAAACCAGC 43

RESULT 13
TS62671
ID T62671 standard; DNA; 45 BP.
AC T62671;
DT 22-MAY-1997 (first entry)
DE 3-hydroxysteroid oxidase gene probe N2.
KW Plant; insect; pest; control; insecticide; toxin; ingestion; probe;
```

```
KW 3-hydroxysteroid oxidase; boll weevil; lepidoptera; kill; stunt growth;
KW larvae; gut; ingestion; hybridisation; Streptomyces sp. A19249; ss.
OS Synthetic.
PN US5538662-A.
PD 24-SEP-1996.
PF 23-SEP-1991; 762682.
PR 23-SEP-1991; US-762682.
PR 04-SEP-1992; US-937195.
PR 28-JUN-1993; US-083948.
PR 24-FEB-1995; US-393785.
PR 07-JUN-1995; US-475694.
PA (MONS ) MONSANTO CO.
PI Corbin DR, Greenplate JT, Jennings MG, Purcell JP,
PI Sammons RD;
DR WPI: 96-442362/44.
PT Genetically transformed plant-colonising microorganism contg.
PT 3-hydroxy;steroid oxidase gene - used for the control of insects
PT esp. lepidopteran and boll weevil.
PS Disclosure; Column 13-14; 18pp; English.
CC Plant insect pests can be controlled by applying to the plant environment
CC or plant seed an insecticidally effective amt. of a protein toxin for
CC ingestion by the insect. In particular, a plant-colonising microorganism
CC genetically engineered to contain and express a heterologous gene
CC comprising a DNA sequence encoding 3-hydroxysteroid oxidase (3-HO) is
CC applied to the plant environment or seed. The protein toxin controls
CC insect infestation, especially boll weevil or lepidopteran. The protein
CC is believed to kill or stunt growth of boll weevil larvae and stunt the
CC growth of lepidopteran larvae by some action in the gut after ingestion.
CC T62671-73 are hybridisation probes used to identify the 3-HO gene in
CC Streptomyces sp. A19249. They were designed based on peptides from the
CC 3-HO protein.
SQ Sequence 45 BP; 8 A; 14 C; 15 G; 8 T;

Query Match 43.7%; Score 16.6; DB 1; Length 45;
Best Local Similarity 71.0%; Pred. No. 3e+02;
Matches 22; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 atgctgacgttgaccatttgaccaca 32
   1:|||||:|:|:|:|:|:|:|:|:|
DB 13 ATGCTGAGATGGCCAGCTGTGAAACCAGC 43

RESULT 14
V11816
ID V11816 standard; DNA; 45 BP.
AC V11816;
DT 11-AUG-1998 (first entry)
DE Streptomyces sp. strain A19249 3-hydroxysteroid oxidase N2 probe.
KW 3-hydroxysteroid oxidase; lepidoptera; plant infestation; CryIA protein;
KW Bacillus thuringiensis; insecticide; cholesterol; crystal protein;
KW synergism; mortality; stunting; larvae; probe; ss.
OS Synthetic.
PN US5763245-A.
PD 03-JUN-1998.
PF 10-SEP-1996; 712057.
PR 10-SEP-1996; US-712057.
PR 23-SEP-1991; US-762682.
PR 04-SEP-1992; US-937195.
PR 28-JUN-1993; US-083948.
PR 24-FEB-1995; US-393785.
PR 07-JUN-1995; US-475694.
PA (MONS ) MONSANTO CO.
PI Corbin DR, Greenplate JT, Pershing JC, Purcell JP,
PI WPI: 98-347327/30.
DR P-PsDB; W59120.
PT Insect control on plants - comprising administration of Bacillus
PT thuringiensis CryIA protein in combination with 3-hydroxy;steroid
PT oxidase
PS Disclosure; Column 29; 25pp; English.
CC This sequence is a hybridisation probe which is used in a method for
CC controlling lepidopteran infestation of plants. This method
```

CC Involves the construction of a recombinant expression vector encoding a
 CC 3-hydroxysteroid oxidase and a Bacillus thuringiensis CryIA protein
 CC each sequence being operatively linked to a promoter to express the
 CC nucleic acid sequences separately or in tandem. This construct can be
 CC used to produce a genetically transformed plant that produces an
 CC insecticidally effective amount of a CryIA protein and a 3-hydroxysteroid
 CC oxidase. combinations of CryIA crystal proteins and cholesterol oxidase
 CC have synergistically enhanced effects on mortality and stunting of
 CC lepidopteran larvae.
 SQ Sequence 45 BP; 8 A; 14 C; 15 G; 8 T;

Query Match 43.7%; Score 16.6; DB 1; Length 45;
 Best Local Similarity 71.0%; Pred. No. 3e+02;
 Matches 22; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 2 atgtcagcttgaccatttgccaccac 32
 ||||| | | | | | | | | | |
 DB 13 ATGCTGAGATGGCCGCTGGAACGAGC 43

RESULT 15

XI9304/c
 ID XI9304 standard; DNA; 41 BP.
 AC XI9304;
 DT 17-MAY-1999 (first entry)
 DE Human granulocyte colony-stimulating factor sense primer oligo 8.
 KW Human; granulocyte colony-stimulating factor; G-CSF; hg-CSF;
 KW haematopoiesis disorder; primer; ss.
 OS Synthetic.
 OS Homo sapiens.
 PN W09853072-A1.
 PD 26-NOV-1998.
 PF 21-MAY-1998; KR0125.
 PR 22-MAY-1997; KR-020054.
 PA (JELI-) JELI PHARM CO LTD.
 PA (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOGY.
 PI Choi B, Jun H, Lee Y, Sohn M;
 PI WPI; 99-131691/11.
 PT New human granulocyte colony-stimulating factor (hg-CSF) containing
 PT a peptide at the N-terminus - useful in the treatment of
 PT haematopoiesis disorders
 PS Example 1; Fig 1; 19pp; English.
 CC The present invention describes the N-terminal peptide of human
 CC granulocyte colony-stimulating factor (hg-CSF). The present invention
 CC also describes: (1) recombinant plasmid pYHM-G-CSF containing the cDNA
 CC for the peptide; and (2) E. coli BL21 (pYHM-G-CSF) (KCTC 0477BP)
 CC transformed with the plasmid. The plasmid and E. coli strain are used
 CC to produce high yields of highly purified hg-CSF, which when refolded
 CC has comparable biological activity of G-CSF. hg-CSF is used to treat
 CC haematopoiesis disorders. The method of producing hg-CSF is less time-
 CC consuming and, therefore, more economical than previous methods because
 CC the protein is expressed in high yields without the need for further
 CC processing steps. X19290 to X19315 represent primer used in the
 CC synthesis of hg-CSF from an example of the present invention.
 SQ Sequence 41 BP; 6 A; 14 C; 11 G; 10 T;

Query Match 43.2%; Score 16.4; DB 1; Length 41;
 Best Local Similarity 67.6%; Pred. No. 3.5e+02;
 Matches 23; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 4 gctgcagcttgaccatttgccaccacatag 37
 ||||| | | | | | | | | | |
 DB 39 GCTGCAGTGTGTCCAAAGTGGAGCCCACTCCGG 6

Search completed: June 4, 2000, 16:24:21
 Job time: 28917 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:32 ; Search time 4521.53 Seconds
(without alignments)
34.064 Million cell updates/sec

Title: US-09-164-714-18

Perfect score: 38
Sequence: 1 catgctgcagcttgaccacatttggcaccacacataggg 38

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 08
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
C 1	18	47.4	67	34	AA474067	AA474067 ves4e02.r
C 2	17.4	45.8	85	35	AA596087	AA596087 nbs5h1.s
C 3	17	44.7	58	28	AA107577	AA107577 mp05a11.r
C 4	17	44.7	97	35	AA580891	AA580891 nc81g07.s
C 5	16.8	44.2	92	23	R85175	R85175 y043d03.r1
C 6	16	42.1	61	63	A1938341	A1938341 sc43e05.y
C 7	16	42.1	67	34	AA465907	AA465907 ves8a04.s
C 8	16	42.1	79	29	AA465762	AA465762 ms60f11.r
C 9	16	42.1	97	25	W05206	W05206 za42h07.r1
C 10	15.8	41.6	76	40	AA930067	AA930067 vg58c03.r
C 11	15.8	41.6	84	30	AA236006	AA236006 zs05a07.s
C 12	15.8	41.6	87	84	B32761	B32761 HS-1015-B2-
C 13	15.8	41.6	90	29	AA155501	AA155501 mr90a11.r
C 14	15.8	41.6	91	47	A1475433	A1475433 l183d04.x
C 15	15.8	41.6	100	27	AA016655	AA016655 mg89h09.r
C 16	15.6	41.1	66	30	AA209802	AA209802 mo79f07.r
C 17	15.6	41.1	100	71	AA158144	AA158144 za62a08.s
C 18	15.4	40.5	46	43	A1209047	A1209047 q928e02.x
C 19	15.4	40.5	77	36	AA606582	AA606582 vm82d10.r
C 20	15.4	40.5	87	81	AA410999	AA410999 fh09g01.y
C 21	15.4	40.5	100	27	AA033513	AA033513 HPLA CCL5
C 22	15.2	40.0	66	23	H55131	H55131 CHR220070 C
C 23	15.2	40.0	80	64	AA072933	AA072933 xa38c12.x
C 24	15.2	40.0	85	36	AA656561	AA656561 vs22c01.r
C 25	15.2	40.0	97	36	AA607212	AA607212 vm84g07.r
C 26	15	39.5	46	46	A1439347	A1439347 t154f06.x
C 27	15	39.5	66	33	AA451186	AA451186 vf88g01.r
C 28	15	39.5	73	51	A1739189	A1739189 w127c11.x
C 29	15	39.5	76	48	A1558686	A1558686 tq31c05.x
C 30	15	39.5	79	40	AA929579	AA929579 vz05b08.r
C 31	15	39.5	82	51	A1745436	A1745436 tr04g11.x
C 32	15	39.5	86	36	AA637197	AA637197 vr31d09.r
C 33	15	39.5	89	23	H44598	H44598 y017c09.r1
C 34	15	39.5	93	39	AA862333	AA862333 og94a05.s
C 35	15	39.5	93	44	A1282700	A1282700 qt83h07.x
C 36	15	39.5	97	33	AA386446	AA386446 vc76h03.r
C 37	15	39.5	97	37	AA714476	AA714476 nw06f12.s
C 38	15	39.5	98	41	AU008426	AU008426 AU008426
C 39	15	39.5	99	70	AA145122	AA145122 ga29e08.y
C 40	15	39.5	100	42	A1159303	A1159303 vz85f05.r
C 41	14.8	38.9	59	35	AA566933	AA566933 987 L0b1o
C 42	14.8	38.9	59	35	AA588419	AA588419 no22a01.s
C 43	14.8	38.9	68	37	AA667085	AA667085 vr87c12.s
C 44	14.8	38.9	82	40	AA936871	AA936871 o151d12.s
C 45	14.8	38.9	85	40	AA917332	AA917332 on45c03.s

ALIGNMENTS

RESULT 1
LOCUS AA474067 67 bp mRNA EST 18-JUN-1997
DEFINITION ves4e02.r1 Beddington mouse embryonic region Mus musculus cDNA
clone IMAGE:821978 5' similar to TR:G476095 G476095 B4-2 PROTEIN.
/, mRNA sequence.

ACCESSION AA474067
VERSION AA474067
KEYWORDS EST.
ORGANISM Mus musculus
SOURCE house mouse.

REFERENCE 1 (bases 1 to 67)
AUTHORS Matra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)

COMMENT

On May 9, 1995 this sequence version replaced g1:803068.
Contact: Maira M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:490258

FEATURES

source

1. 67
/organism="Mus musculus"
/strain="C57BL6 x DBA"
/db_xref="taxon:10090"
/clone="IMAGE:821978"
/clone_lib="Beddington mouse embryonic region"
/sex="pooled"
/tissue_type="embryo"
/dev_stage="7.5dpc"
/lab_host="DH12S"
/note="Organ: whole embryo; Vector: PCMV-SPORT; Site:1:
SalI; Site:2: NotI; Cloned unidirectionally. Primer:
Oligo dt. Gastrulating embryos were collected at 7.5dpc
from C57BL6 x DBA matings, excluding embryos that had
developed head folds and all extraembryonic tissues.
Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
Referenced in Development 121, 2479-2489 (1995)"

BASE COUNT
ORIGIN

19 a 23 c 13 g 12 t

Query Match 47.4%; Score 18; DB 34; Length 67;
Best Local Similarity 70.6%; Pred. No. 1e+03;
Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 3 tgcgcagcttgaccacatttgaccacacatag 36
db 37 TGAATGGCGTCCATTAATTGGCTCCACACATAG 4

RESULT 2

LOCUS AA596087 85 bp mRNA EST 26-SEP-1997
DEFINITION nbs5h1.s1 NCI-CGAP_Br2 Homo sapiens cDNA clone IMAGE:1098021 3'
similar to gb:U19872 AH RECEPTOR PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AA596087
VERSION AA596087
KEYWORDS EST.
ORGANISM human.

REFERENCE 1 (bases 1 to 85)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE Robert Strausberg@nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

CONTACT: Robert Strausberg, Ph.D.
On Apr 14, 1993 this sequence version replaced g1:693426.
Tel: (301) 496-1550

JOURNAL

COMMENT

Query Match	44.7%	Score 17;	DB 35;	Length 97;
Best Local Similarity	69.7%	Pred. No. 2.9e+03;		
Matches 23; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0

```

QY      2 atcgtcagcttgaccaatttggcaccacat 34
          | | | | | | | | | | | | | | | |
Db      90 AGGCTGCAATGCAGCAGATATTGCACCCACAT 58

```

DEFINITION	LOCUS	RESULT	5
yo43403.r1 Soares adult brain N2d4HB55Y Homo sapiens cDNA clone	R85175	92 bp	EST 14-AUG-1995

ORGANISM	Homo sapiens

BASE COUNT	18 a	24 c	26 g	18 t	6 others
ORIGIN	basal ganglia, thalamus, cerebellum, midbrain, pons and medulla."				

Query Match	44.28;	Score 16.8;	DB 23;	Length 92;
Best Local Similarity	64.98;	Pred. No. 3.5e+03;		
Matches 24; Conservative	0;	Mismatches 13;	Indels 0;	Gaps 0

Qy 1 catgctgcagctcttgaaccaattttgtgacccaacatag 37
|| ||||| ||| |||||
Db 21 CACCCTCGACGCTTGCCCACTNCGGCAACCAAGATTTCG 57

RESULT	6
AI938341/c	
LOCUS	AI938341
DEFINITION	61 bp mRNA
	sc43e05.y1 Gm-c1014 glycine max cDNA clone
	EST
	02-AUG-1999
	GENOME SYSTEMS CLONE ID

SOURCE
ORGANISM
soybean.
glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
glycine.

FEATURES	Location/Qualifiers
source	1. .92

FEATURES
source

were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker and Dr.

BASE COUNT	23 a	12 c	17 g
ORIGIN			

Query Match	42.1%	Score 16;	DB 63;	Length 61;
Best Local Similarity	79.2%	Pred. No. 6.6e+03;		
Matches 19; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;

```

QY      3  tgcctgcagcttgaccaatttggc 26
          | ||||| ||||| ||||| ||
Db      49  TCCTGCACACTTGGCCAACTCTTGC 26

```

RESULT	7			
LOCUS	AA465907			
DEFINITION	AA465907	67 bp	mrna	EST
IMAGE:	ve99a44.s1	Knowles	Solter mouse 2 cell	Mus musculus cdna clone
	IMAGE:833358	5'	mrna sequence.	

FEATURES	Location/Qualifiers
source	1. .67

BASE COUNT
ORIGIN

Query Match	42.1%	Score 16;	DB 34;	Length 67;
Best Local Similarity	68.8%;	Pred. No. 6.8e+03;		
Matches 22;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0

QY 1 catgctgacgtctgaccaatttggcaccac 32
||| ||||| ||| ||||| |||
Db 13 CAGCTTCCAGCTTGCGCTTGTGTGGCACCAC 44

RESULT	8
AA165762	79 bp mRNA
LOCUS	EST
DEFINITION	12-FEB-1997
	mus0011.1 Stratagene mouse embryonic carcinoma (#937317) MUS
	musculi cDNA IMAGE:61581 5' similar to TR:E93245 E93245 ETN
	INSERT IN THE FAS APOPTOSIS GENE OF MRL-IPR/IPR. [1] ; mRNA
	sequence.

FEATURES

BASE COUNT
ORIGIN

Query Match	42.1%	Score 16:	DB 29:	Length 79:
Best Local Similarity	68.8%	Pred. No.	7.1e+03:	
Matches 22, Conservative	0:	Mismatches 10:	Indels 0:	Gaps 0:

```

Qy      7  gcagcttgaccaatttggcaccacaataggg 38
          |||  |  |||||  |||  |  |||
Db      25  GCGGCGCACATTTTGGCGCCAGACTGG 56

```

RESULT	9
W05206/c	
LOCUS	W05206
DEFINITION	97 bp mRNA EST 23-APR-1996 zaa2h07.r1 Soares fetal liver spleen INELIS Homo sapiens cDNA clone IMAGE:295261 5' similar to PIR:S45038 S45038 CABP1 protein - rat ;,

ACCESSION mRNA sequence.
W05206
VERSION W05206.1 GI:1277938
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 97)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevisan,E., Waterston,R., Williamson,A., Woldmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT On May 8, 1995 this sequence version replaced gi:800180.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watsn.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: mob.REGA-ET
High quality sequence stop: 48.
Location/Qualifiers
1..97
/organism="Homo sapiens"
/db_xref="GDB:1240182"
/db_xref="taxon:9606"
/clone="IMAGE:295261"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site.1: Pac I; Site.2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5', AACGAGAGATTAATTAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."
BASE COUNT 36 a 15 c 23 g 22 t 1 others
ORIGIN
Query Match 42.1%; Score 16; DB 25; Length 97;
Best Local Similarity 68.8%; Pred. No. 7.5e+03;
Matches 22; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 3 tgcgcagcttgaccatttgcaccacat 34
111 11 11 1111 111 11 11 11
DB 73 TGCATCACTGCACCAACTTGCACACACTT 42
RESULT 10
AA930067 76 bp mRNA EST 23-APR-1998
LOCUS AA930067
DEFINITION vq58c03.r1 Barstead mouse proximal colon MRLRB6 Mus musculus cDNA
clone IMAGE:1106500 5' similar to SM:ATPM_MOUSE_P56384 ATP SYNTHASE
LIPID-BINDING PROTEIN P3 PRECURSOR; mRNA sequence.
ACCESSION AA930067
VERSION AA930067.1 GI:3079660
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 76)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMT Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2285643.
Contact: Marra M/Mouse EST Project
WashU-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsn.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:604668
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28m3 rev2 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..76
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1106500"
/clone_lib="Barstead mouse proximal colon MRLRB6"
/dev_stage="7 day juvenile"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site.1: EcoRI; Site.2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTCAGATCTGAGTGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTT
3]], double-stranded cDNA was ligated to Eco RI adaptors
[AATTCGATCTCTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."
BASE COUNT 18 a 22 c 16 g 20 t
ORIGIN
Query Match 41.6%; Score 15.8; DB 40; Length 76;
Best Local Similarity 74.1%; Pred. No. 8.5e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 gctgcagcttgaccatttgcaccac 30
1111 11111111 111 11
DB 29 GTTGCAATATAGACCAATTTCTGCATCA 55
RESULT 11
AA236006 84 bp mRNA EST 13-AUG-1997
LOCUS AA236006
DEFINITION zs05a07.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:684276 3',
mRNA sequence.
ACCESSION AA236006
VERSION AA236006.1 GI:1860444
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 84)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1290618.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov

Query Match 41.6%; Score 15.8; DB 29; Length 90;
 Best Local Similarity 65.7%; Pred. No. 9e+03;
 Matches 23; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 3 tgcgcagcttgaccattggcaccacatagg 37
 ||||| ||||| ||||| ||||| |||||
 Db 81 tgcgccttattgaccattttaaaccctaaatpaaag 47

RESULT 14
 A1475433 91 bp mRNA EST 09-MAR-1999
 LOCUS C183d04.x1 NCI_CGAP_Col4 Homo sapiens cDNA clone IMAGE:2153671 3',
 DEFINITION mRNA sequence.
 A1475433
 VERSION A1475433.1 GI:4328478
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 91)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On Jun 22, 1998 this sequence version replaced gi:3247120.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bbrp/image/image.html

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Seq primer: -40UP from Gibco
 High quality sequence stop: 85.
 Location/Qualifiers
 1..91
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2153671"
 /clone_lib="NCI_CGAP_Col4"
 /tissue_type="moderately-differentiated adenocarcinoma"
 /lab_host="DH10B"
 /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.7 kb. Life Technologies catalog #: 11531-019"

BASE COUNT 8 a 5 c 20 g 58 t
 ORIGIN

Query Match 41.6%; Score 15.8; DB 47; Length 91;
 Best Local Similarity 65.7%; Pred. No. 9e+03;
 Matches 23; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 4 gctgcagcttgaccattggcaccacatagg 38
 ||||| ||||| ||||| ||||| |||||
 Db 49 GCGGGGCTGTGTTTATTTTGGCCCAAAAAGG 83

RESULT 15
 AA016655 100 bp mRNA EST 02-AUG-1996
 LOCUS AA016655
 DEFINITION mg89h09.tl Soares mouse embryo NDM13.5 14.5 Mus musculus cDNA
 clone IMAGE:440225 5' similar to SW:ATPM_HUMAN Q06055 ATP SYNTHASE
 LIPID-BINDING PROTEIN P2 PRECURSOR ;, mRNA sequence.
 ACCESSION AA016655

VERSION AA016655.1 GI:1478885
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 100)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HMIT Mouse EST Project
 Unpublished (1996)
 On Apr 14, 1993 this sequence version replaced gi:693329.
 Contact: Marra M/Mouse EST Project
 WashU-HMIT Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:265561
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers
 1..100
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:440225"
 /clone_lib="Soares mouse embryo NDM13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTCACATCTGAGTGGAGCGCGCGAATAATTTTTTTTTTTTTTTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2]; double-stranded cDNA was ligated to
 Eco RI adaptors (pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT7T3 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."

BASE COUNT 20 a 33 c 23 g 24 t
 ORIGIN

Query Match 41.6%; Score 15.8; DB 27; Length 100;
 Best Local Similarity 74.1%; Pred. No. 9.2e+03;
 Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 gctgcagcttgaccattggcaccac 30
 ||||| ||||| ||||| ||||| |||||
 Db 56 GTTGCAATATAGACAATTCCTGCATCA 82

Search completed: June 4, 2000, 13:53:34
 Job time: 20958 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:47 ; Search time 244.64 Seconds
(without alignments)
20.191 Million cell updates/sec

Title: US-09-164-714-18

Perfect score: 38

Sequence: 1 catgctgcagcttgaccatttggcaccacaataagag 38

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 segs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/6.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/PCUS.COMB.seq:*
7: /cgn2_6/ptodata/2/1na/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19.6	51.6	96	1	US-07-835-029A-2
2	19.6	51.6	96	1	US-08-049-511-2
3	19.6	51.6	96	1	US-07-934-554-2
4	19.6	51.6	96	6	PCT-US93-07877-2
5	19.6	51.6	96	6	PCT-US94-04151-2
6	17.4	45.8	63	6	PCT-US95-13975-63
7	16.6	43.7	45	1	US-08-083-948-4
8	16.6	43.7	45	1	US-08-393-785-4
9	16.6	43.7	45	1	US-08-475-694-4
10	16.6	43.7	45	2	US-08-712-057-4
11	16.6	43.2	60	1	US-07-734-225A-11
12	16.4	43.2	60	1	US-07-692-995B-11
13	16.4	43.2	60	2	US-08-488-457-11
14	16.4	43.2	60	3	US-08-338-793D-11
15	16.4	43.2	90	3	US-08-431-459A-11
16	16.4	43.2	90	6	PCT-US95-13975-70
17	16.4	42.1	39	1	US-08-361-920-32
18	16.4	42.1	39	1	US-08-479-939-32
19	16.4	42.1	39	2	US-08-483-432-32
20	15.8	41.6	45	3	US-08-750-856A-4
21	15.4	40.5	50	1	US-08-207-901-75
22	15.4	40.5	63	6	PCT-US95-13975-49
23	15.2	40.0	51	1	US-08-207-996-10
24	15.2	40.0	51	4	US-08-760-840A-10
25	15.2	40.0	63	1	US-07-734-225A-12
26	15.2	40.0	63	1	US-07-692-995B-12
27	15.2	40.0	63	2	US-08-488-457-12

28	15.2	40.0	63	3	US-08-338-793D-12	Sequence 12, Appl
29	15.2	40.0	63	3	US-08-431-459A-12	Sequence 12, Appl
30	15.2	40.0	80	1	US-07-971-160-50	Sequence 50, Appl
31	15.2	40.0	80	2	US-08-336-241-50	Sequence 50, Appl
32	15.2	40.0	80	3	US-08-465-273-50	Sequence 50, Appl
33	15.2	40.0	80	4	US-09-119-024-50	Sequence 50, Appl
34	15.2	40.0	80	4	US-08-417-226-50	Sequence 50, Appl
35	15.2	40.0	85	1	US-07-741-931-5	Sequence 5, Appl
36	15.2	40.0	85	1	US-07-937-132A-5	Sequence 5, Appl
37	15.2	40.0	85	1	US-07-937-132A-6	Sequence 6, Appl
38	15.2	40.0	86	1	US-07-741-931-7	Sequence 7, Appl
39	15.2	40.0	86	1	US-07-937-132A-7	Sequence 7, Appl
40	15.2	39.5	34	4	US-08-459-135A-2	Sequence 2, Appl
41	14.8	38.9	98	3	US-08-731-272A-4	Sequence 4, Appl
42	14.4	37.9	36	1	US-07-961-884A-21	Sequence 21, Appl
43	14.4	37.9	36	2	US-08-428-941-21	Sequence 21, Appl
44	14.4	37.9	63	6	PCT-US95-13975-15	Sequence 15, Appl
45	14.4	37.9	63	6	PCT-US95-13975-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-07-835-029A-2/C
Sequence 2, Application US/07835029A
Patent No. 5284664
GENERAL INFORMATION:
APPLICANT: WAGLE, SUDHAKAR S
APPLICANT: STEINBACH, THOMAS
APPLICANT: LAWYER, CARL H
APPLICANT: HERMANN, WILLIAM J
APPLICANT: GAWISH, ALI ABDEL SALAM
TITLE OF INVENTION: Method of Treating Presenile or Senile
TITLE OF INVENTION: Dementia
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut
STREET: 100 South Wacker Drive, Hartford Plaza
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-4002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/835,029A
FILING DATE: 19920205
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fentress, Susan B
REGISTRATION NUMBER: 31,327
REFERENCE/DOCKET NUMBER: 91179A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/456-8000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-07-835-029A-2

Query Match 51.6%; Score 19.6; DB 1; Length 96;
Best Local Similarity 73.5%; Pred. No. 11;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 2 atgctgcagcttgaccatttggcaccacaata 35


```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07877
; FILING DATE: 19930823
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pentress, Susan B
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92011A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/456-8000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; PCT-US93-07877-2

Query Match          51.6%; Score 19.6; DB 6; Length 96;
Best Local Similarity 73.5%; Pred. No. 11;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 atgtcgacgttgaccatttggcaccacata 35
Db 79 AAGCTGCAGCTTGACGTATTATTGCACTTGATGATA 46

RESULT 5
PCT-US94-04151-2/c
; Sequence 2, Application PC/TUS9404151
; GENERAL INFORMATION:
; APPLICANT: MAGLE, SUDHAKAR S
; APPLICANT: STEINBACH, THOMAS
; APPLICANT: LAWYER, CARL H
; APPLICANT: HERMANN, WILLIAM J
; APPLICANT: GAMISH, ALI ABDEL SALAM
; TITLE OF INVENTION: METHOD OF TREATING HEPATITIS B
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TILTON, FALTON, LUNGWUS & CHESTNUT
; STREET: 100 SOUTH WACKER DRIVE, SUITE 960
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04151
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/7780,084
; FILING DATE: 15-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/804,844
; FILING DATE: 04-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,267
; FILING DATE: 11-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/228,364
; FILING DATE: 04-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: PENTRESS, SUSAN B
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92008A
; PCT-US94-04151-2
```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/456-8000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; PCT-US94-04151-2

Query Match          51.6%; Score 19.6; DB 6; Length 96;
Best Local Similarity 73.5%; Pred. No. 11;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 atgtcgacgttgaccatttggcaccacata 35
Db 79 AAGCTGCAGCTTGACGTATTATTGCACTTGATGATA 46

RESULT 6
PCT-US95-13975-63
; Sequence 63, Application PC/TUS9513975
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: McDONELL, Michael W.
; TITLE OF INVENTION: Recombinant feline Herpes virus
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM 330 466 DX2
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13975
; FILING DATE: 26-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/329,883
; FILING DATE: 26-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39118-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0525
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudorabies virus; Feline herpesvirus
; PCT-US95-13975-63

Query Match          45.8%; Score 17.4; DB 6; Length 63;
Best Local Similarity 77.8%; Pred. No. 79;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
; PCT-US95-13975-63
```

OY 4 gctgcagcttgaccatttgcacca 30
||||| | | | | | | | | |
Db 36 GCTGCAGGTGACGAGTCTACACCA 62

RESULT 7

US-08-083-948-4
; Sequence 4, Application US/08083948
; Patent No. 5518908
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Greenplate, Michael G.
; APPLICANT: Jennings, John P.
; APPLICANT: Purcell, John P.
; APPLICANT: Sammons, Robert D.
; TITLE OF INVENTION: Method of Controlling Insects
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
; STREET: 700 Chesterfield Parkway No. 5518908th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/083,948
; FILING DATE: 19930628
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/762682
; FILING DATE: 23-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937195
; FILING DATE: 09-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(10631)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-7286
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-083-948-4

Query Match 43.7%; Score 16.6; DB 1; Length 45;
Best Local Similarity 71.0%; Pred. No. 1.5e+02;
Matches 22; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 2 atgctgcagcttgaccatttgcacca 32
||||| | | | | | | | | |
Db 13 ATGCTGAGATGGCCAGCTGTGAACACG 43

RESULT 8

US-08-393-785-4
; Sequence 4, Application US/08393785
; Patent No. 5554369
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Greenplate, John T.

APPLICANT: Jennings, Michael G.
APPLICANT: Purcell, John P.
APPLICANT: Sammons, Robert D.
TITLE OF INVENTION: Method of Controlling Insects
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
STREET: 700 Chesterfield Parkway No. 5554369th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,785
FILING DATE: 24-FEB-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/083,948
FILING DATE: 28-JUN-1993
APPLICATION NUMBER: US 07/762682
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937195
FILING DATE: 09-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10631)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-7286
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-393-785-4

Query Match 43.7%; Score 16.6; DB 1; Length 45;
Best Local Similarity 71.0%; Pred. No. 1.5e+02;
Matches 22; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 2 atgctgcagcttgaccatttgcacca 32
||||| | | | | | | | | |
Db 13 ATGCTGAGATGGCCAGCTGTGAACACG 43

RESULT 9

US-08-475-694-4
; Sequence 4, Application US/08475694
; Patent No. 5558862
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Greenplate, John T.
; APPLICANT: Jennings, Michael G.
; APPLICANT: Purcell, John P.
; APPLICANT: Sammons, Robert D.
; TITLE OF INVENTION: Method of Controlling Insects
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
; STREET: 700 Chesterfield Parkway No. 5558862th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA


```
; FILING DATE: 23-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9018418.5
; FILING DATE: 23-AUG-1990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)861-3000
; TELEFAX: (202)822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-734-225A-11

Query Match          43.2%; Score 16.4; DB 1; Length 60;
Best Local Similarity 67.6%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 4 gctgcagcttgacccaatttgaccacacatag 37
    |||||  || |||  || |||||  |||
Db 34 GCTGCAGTGTGTCACAGGTGGCCCAATTTCAGG 1

RESULT 12
US-07-692-995B-11/c
; Sequence 11, Application US/07692995B
; Patent No. 5416195
; GENERAL INFORMATION:
; APPLICANT: Camble, Roger
; APPLICANT: Heather, Carr
; APPLICANT: Timms, David
; APPLICANT: Wilkinson, Anthony J.
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: Eleventh Floor, 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/692,995B
; FILING DATE: 19910429
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9009623.1
; FILING DATE: 30-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9013773.8
; FILING DATE: 20-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9016215.7
; FILING DATE: 24-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9102799.5
; FILING DATE: 11-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: NUCLEIC ACID
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-692-995B-11

Query Match          43.2%; Score 16.4; DB 1; Length 60;
Best Local Similarity 67.6%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 4 gctgcagcttgacccaatttgaccacacatag 37
    |||||  || |||  || |||||  |||
Db 34 GCTGCAGTGTGTCACAGGTGGCCCAATTTCAGG 1

RESULT 13
US-08-488-457-11/c
; Sequence 11, Application US/08488457
; Patent No. 5773581
; GENERAL INFORMATION:
; APPLICANT: Camble, Roger
; APPLICANT: Timms, David
; APPLICANT: Wilkinson, Anthony J.
; TITLE OF INVENTION: CONTINUOUS RELEASE PHARMACEUTICAL
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: Eleventh Floor, 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,457
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/155,327
; FILING DATE: 22-NOV-1993
; APPLICATION NUMBER: US 07/734,225
; FILING DATE: 22-JUL-1991
; APPLICATION NUMBER: GB 9016138.1
; FILING DATE: 23-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9018414.4
; FILING DATE: 23-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9018415.1
; FILING DATE: 23-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9018416.9
; FILING DATE: 23-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9018417.1
; FILING DATE: 23-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9018418.5
; FILING DATE: 23-AUG-1990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)861-3000
; TELEFAX: (202)822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

US-08-488-457-11

Query Match 43.2%; Score 16.4; DB 2; Length 60;
Best Local Similarity 67.6%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 4 gctgcagcttgaccatttggcaccacatagg 37
||||| ||| | ||| ||| |||
DB 34 GCTGCAGTGTGTCCAGGTTGGGCCCAATTCAGG 1

RESULT 14.

US-08-338-793D-11/c
; Sequence 11, Application US/08338793D
; Patent No. 5840521
; GENERAL INFORMATION:
; APPLICANT: Barth, Peter Thomas
; TITLE OF INVENTION: VECTOR
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY CUSHMAN
; ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM PC/XT/AT Compatibles
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word or ASCII editors
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/338,793D
; FILING DATE: 08-No. 5840521-94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/842,081
; FILING DATE: 26-Feb-92
; CLASSIFICATION: 435
; APPLICATION NUMBER: 9104017.0
; FILING DATE: 26-Feb-91
; APPLICATION NUMBER: 9109188.4
; FILING DATE: 29-Apr-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Kokulis, Paul N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: DJB/9901/215431/TGW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-338-793D-11

Query Match 43.2%; Score 16.4; DB 3; Length 60;
Best Local Similarity 67.6%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 4 gctgcagcttgaccatttggcaccacatagg 37
||||| ||| | ||| ||| |||
DB 34 GCTGCAGTGTGTCCAGGTTGGGCCCAATTCAGG 1

RESULT 15

US-08-431-459A-11/c
; Sequence 11, Application US/08431459A
; Patent No. 5840543

; GENERAL INFORMATION:
; APPLICANT: Hockney, Robert C.
; APPLICANT: Kara, Bhupendra V.
; TITLE OF INVENTION: FERMENTATION PROCESS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44m diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,459A
; FILING DATE: 01-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,676
; FILING DATE: 22-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/661,306
; FILING DATE: 27-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9004390.2
; FILING DATE: 27-FEB-1990
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-431-459A-11

Query Match 43.2%; Score 16.4; DB 3; Length 60;
Best Local Similarity 67.6%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 4 gctgcagcttgaccatttggcaccacatagg 37
||||| ||| | ||| ||| |||
DB 34 GCTGCAGTGTGTCCAGGTTGGGCCCAATTCAGG 1

Search completed: June 4, 2000, 16:09:48
Job time: 28067 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:44 ; Search time 1236.38 Seconds
(without alignments)
-31.472 Million cell updates/sec

Title: US-09-164-714-19
Perfect score: 40
Sequence: 1 cactctgcagtagacgcacagcaatcaacgtagaaca 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenDBml:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_hg1:*
33: gb_hg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_hg3:*
42: gb_hg4:*
43: gb_hg5:*
44: gb_hg6:*

45: gb_hg7:*
46: em_hg1:*
47: em_hg2:*
48: em_hg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_hg8:*
53: gb_hg9:*
54: gb_hg10:*
55: gb_hg11:*
56: gb_hg12:*
57: gb_hg13:*
58: gb_hg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19.2	48.0	54	5 A57959	A57959 Sequence 25
2	18.2	45.5	54	5 A57952	A57952 Sequence 18
3	18.2	45.5	54	5 A57960	A57960 Sequence 26
4	17.2	43.0	54	5 A57944	A57944 Sequence 10
5	17.2	43.0	54	5 A57945	A57945 Sequence 11
6	17.2	43.0	54	5 A57946	A57946 Sequence 12
7	17.2	43.0	54	5 A57947	A57947 Sequence 13
8	17.2	43.0	54	5 A57948	A57948 Sequence 14
9	17.2	43.0	54	5 A57949	A57949 Sequence 15
10	17.2	43.0	54	5 A57951	A57951 Sequence 17
11	17.2	43.0	54	5 A57954	A57954 Sequence 20
12	17.2	43.0	54	5 A57957	A57957 Sequence 22
13	17.2	43.0	54	5 A57961	A57961 Sequence 27
14	17.2	43.0	54	5 E05823	E05823 DNA encodin
15	16.8	42.0	54	5 A57950	A57950 Sequence 16
16	16.8	42.0	54	5 A57955	A57955 Sequence 21
17	16.8	42.0	84	16 HCVHR31	299037 Hepatitis C
18	16.8	42.0	84	16 HCVHR33	299039 Hepatitis C
19	16.8	42.0	9	D78279S17	D78295 Homo sapien
20	16.8	42.0	54	5 A57958	A57958 Sequence 24
21	16.2	40.5	57	5 I70194	I70194 Sequence 9
22	16.2	40.5	69	5 AR035210	AR035210 Sequence
23	16	40.0	69	5 AR035215	AR035215 Sequence
24	16	40.0	35	5 A21055	A21055 Synthetic o
25	15.8	39.5	74	12 RATPAM25	U52661 Rattus norv
26	15.8	39.5	96	9 HS001820	A0001820 Homo sapl
27	15.8	39.5	97	5 I91489	I91489 Sequence 23
28	15.6	39.5	54	5 A57943	A57943 Sequence 9
29	15.6	39.0	54	5 A57953	A57953 Sequence 19
30	15.6	39.0	54	5 A57965	A57965 Sequence 31
31	15.6	39.0	58	5 E01587	E01587 DNA sequenc
32	15.6	39.0	58	5 E01745	E01745 DNA sequenc
33	15.6	39.0	61	5 E01585	E01585 DNA sequenc
34	15.6	39.0	84	16 HCVHR10	299016 Hepatitis C
35	15.6	39.0	84	16 HCVHR11	299017 Hepatitis C
36	15.6	39.0	84	16 HCVHR9	299015 Hepatitis C
37	15.6	38.5	65	1 S72520	S72520 flm3-major
38	15.4	38.0	65	5 A21052	A21052 B.pertussis
39	15.2	38.5	81	12 MMR25B9	X53627 M.musculus
40	15.2	38.0	84	16 HCVHR12	299018 Hepatitis C
41	15.2	38.0	84	16 HCVHR13	299019 Hepatitis C
42	15.2	38.0	84	16 HCVHR15	299021 Hepatitis C
43	15.2	38.0	84	16 HCVHR16	299022 Hepatitis C
44	15.2	38.0	84	16 HCVHR17	299023 Hepatitis C
45	15.2	38.0	84	16 HCVHR17	299023 Hepatitis C

ALIGNMENTS

```
RESULT 1
A57959 54 bp DNA PAT 05-MAR-1998
LOCUS Sequence 25 from Patent EP0743364.
DEFINITION A57959
ACCESSION A57959
VERSION A57959.1 GI:3713729
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 54)
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding
JOURNAL fragments and their application as reactives for risk evaluation of
COMMENT HIV-1 mother-foetal transmission
FEATURES Patent: EP 0743364-A 25 20-NOV-1996;
Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 4 c 15 g 5 t
ORIGIN

Query Match 48.0%; Score 19.2; DB 5; Length 54;
Best Local Similarity 75.0%; Pred. No. 9.5e+02;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 9 agtagcgccaagcaaatcaacgctaagaaca 40
| | | | | | | | | | | | | | | | | |
DB 18 ACTAGAGGAAAGACAAACAAAGTAAGAAAA 49

RESULT 2
A57952 54 bp DNA PAT 05-MAR-1998
LOCUS Sequence 18 from Patent EP0743364.
DEFINITION A57952
ACCESSION A57952
VERSION A57952.1 GI:3713722
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 54)
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding
JOURNAL fragments and their application as reactives for risk evaluation of
COMMENT HIV-1 mother-foetal transmission
FEATURES Patent: EP 0743364-A 18 20-NOV-1996;
Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 29 a 5 c 15 g 4 t 1 others
ORIGIN

Query Match 45.5%; Score 18.2; DB 5; Length 54;
Best Local Similarity 74.2%; Pred. No. 2.5e+03;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 10 gtagcgccaagcaaatcaacgctaagaaca 40
| | | | | | | | | | | | | | | | | |
DB 19 GTAGAGGAAAGACAAACAAAGTAAGAAAA 49

RESULT 3
A57960 54 bp DNA PAT 05-MAR-1998
LOCUS Sequence 26 from Patent EP0743364.
DEFINITION A57960
ACCESSION A57960
VERSION A57960.1 GI:3713730
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 54)
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding
JOURNAL fragments and their application as reactives for risk evaluation of
COMMENT HIV-1 mother-foetal transmission
FEATURES Patent: EP 0743364-A 26 20-NOV-1996;
Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 4 c 15 g 5 t
ORIGIN

Query Match 45.5%; Score 18.2; DB 5; Length 54;
Best Local Similarity 74.2%; Pred. No. 2.5e+03;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 10 gtagcgccaagcaaatcaacgctaagaaca 40
| | | | | | | | | | | | | | | | | |
DB 19 GTAGAGGAAAGACAAACAAAGTAAGAAAA 49

RESULT 4
A57944 54 bp DNA PAT 05-MAR-1998
LOCUS Sequence 10 from Patent EP0743364.
DEFINITION A57944
ACCESSION A57944
VERSION A57944.1 GI:3713714
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 54)
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding
JOURNAL fragments and their application as reactives for risk evaluation of
COMMENT HIV-1 mother-foetal transmission
FEATURES Patent: EP 0743364-A 10 20-NOV-1996;
Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 31 a 4 c 13 g 6 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11 tagcgccaagcaaatcaacgctaagaaca 40
| | | | | | | | | | | | | | | | | |
DB 20 TAGAGGAAAGACAAACAAAGTAAGAAAA 49

RESULT 5
A57945 54 bp DNA PAT 05-MAR-1998
LOCUS Sequence 11 from Patent EP0743364.
ACCESSION A57945
```

```
LOCUS A57960 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 26 from Patent EP0743364.
ACCESSION A57960
VERSION A57960.1 GI:3713730
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 54)
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding
JOURNAL fragments and their application as reactives for risk evaluation of
COMMENT HIV-1 mother-foetal transmission
FEATURES Patent: EP 0743364-A 26 20-NOV-1996;
Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 4 c 15 g 5 t
ORIGIN

Query Match 45.5%; Score 18.2; DB 5; Length 54;
Best Local Similarity 74.2%; Pred. No. 2.5e+03;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 10 gtagcgccaagcaaatcaacgctaagaaca 40
| | | | | | | | | | | | | | | | | |
DB 19 GTAGAGGAAAGACAAACAAAGTAAGAAAA 49

RESULT 4
A57944 54 bp DNA PAT 05-MAR-1998
LOCUS Sequence 10 from Patent EP0743364.
DEFINITION A57944
ACCESSION A57944
VERSION A57944.1 GI:3713714
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 54)
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding
JOURNAL fragments and their application as reactives for risk evaluation of
COMMENT HIV-1 mother-foetal transmission
FEATURES Patent: EP 0743364-A 10 20-NOV-1996;
Other publication FR 2734281 961122.
source Location/Qualifiers
1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 31 a 4 c 13 g 6 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11 tagcgccaagcaaatcaacgctaagaaca 40
| | | | | | | | | | | | | | | | | |
DB 20 TAGAGGAAAGACAAACAAAGTAAGAAAA 49

RESULT 5
A57945 54 bp DNA PAT 05-MAR-1998
LOCUS Sequence 11 from Patent EP0743364.
ACCESSION A57945
```


VERSION A57945.1 GI:3713715
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 11 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source
1. 54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 4 c 15 g 5 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagccaagaacaatcaacgctaagaaca 40
||||| ||||||| ||| |||||||
Db 20 TAGAGGAGAGAGCAAAACAAAGTAAGAAAA 49

RESULT 6
LOCUS A57946 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 12 from Patent EP0743364.
ACCESSION A57946
VERSION A57946.1 GI:3713716
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 12 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source
1. 54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 31 a 4 c 14 g 5 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagccaagaacaatcaacgctaagaaca 40
||||| ||||||| ||| |||||||
Db 20 TAGAGGAGAGAGCAAAACAAAGTAAGAAAA 49

RESULT 7
LOCUS A57947 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 13 from Patent EP0743364.
ACCESSION A57947
VERSION A57947.1 GI:3713717
KEYWORDS
SOURCE
unidentified.

ORGANISM
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 13 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source
1. 54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 31 a 4 c 13 g 6 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagccaagaacaatcaacgctaagaaca 40
||||| ||||||| ||| |||||||
Db 20 TAGAGGAGAGAGCAAAACAAAGTAAGAAAA 49

RESULT 8
LOCUS A57948 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 14 from Patent EP0743364.
ACCESSION A57948
VERSION A57948.1 GI:3713718
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 14 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source
1. 54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 32 a 5 c 12 g 5 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagccaagaacaatcaacgctaagaaca 40
||||| ||||||| ||| |||||||
Db 20 TAGAGGAGAGAGCAAAACAAAGTAAGAAAA 49

RESULT 9
LOCUS A57949 54 bp DNA PAT 05-MAR-1998
DEFINITION Sequence 15 from Patent EP0743364.
ACCESSION A57949
VERSION A57949.1 GI:3713719
KEYWORDS
SOURCE
unidentified.
ORGANISM
REFERENCE 1 (bases 1 to 54)

AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 15 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source Location/Qualifiers
1. 54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 5 c 14 g 5 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagccaagcaatcaacggttaagaca 40
||||| ||||||| ||| ||||||| |
Db 20 TAGAGGAGAGCGCAAAACAAAGTAGAGAAA 49

RESULT 10
A57951 54 bp DNA PAT 05-MAR-1998
LOCUS A57951
DEFINITION Sequence 17 from Patent EP0743364.
ACCESSION A57951
VERSION A57951.1 GI:3713721
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 17 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source Location/Qualifiers
1. 54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 31 a 4 c 14 g 5 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagccaagcaatcaacggttaagaca 40
||||| ||||||| ||| ||||||| |
Db 20 TAGAGGAGAGCGCAAAACAAAGTAGAGAAA 49

RESULT 11
A57954 54 bp DNA PAT 05-MAR-1998
LOCUS A57954
DEFINITION Sequence 20 from Patent EP0743364.
ACCESSION A57954
VERSION A57954.1 GI:3713724
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of

JOURNAL HIV-1 mother-foetal transmission
COMMENT Patent: EP 0743364-A 20 20-NOV-1996;
FEATURES COMMISSARIAT ENERGIE ATOMIQUE (FR)
source Other publication FR 2734281 961122.
Location/Qualifiers
1. 54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 4 c 15 g 5 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagccaagcaatcaacggttaagaca 40
||||| ||||||| ||| ||||||| |
Db 20 TAGAGGAGAGCGCAAAACAAAGTAGAGAAA 49

RESULT 12
A57956 54 bp DNA PAT 05-MAR-1998
LOCUS A57956
DEFINITION Sequence 22 from Patent EP0743364.
ACCESSION A57956
VERSION A57956.1 GI:3713726
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 22 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
source Location/Qualifiers
1. 54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 6 c 14 g 4 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagccaagcaatcaacggttaagaca 40
||||| ||||||| ||| ||||||| |
Db 20 TAGAGGAGAGCGCAAAACAAAGTAGAGAAA 49

RESULT 13
A57957 54 bp DNA PAT 05-MAR-1998
LOCUS A57957
DEFINITION Sequence 23 from Patent EP0743364.
ACCESSION A57957
VERSION A57957.1 GI:3713727
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa,R. and Roques,P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-foetal transmission
JOURNAL Patent: EP 0743364-A 23 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)

COMMENT Other publication FR 2734281 961122.
FEATURES Location/Qualifiers
SOURCE 1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 6 c 13 g 5 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11 tagagcccaagcaatcaacggttaagaca 40
||||| ||||| ||| ||||| |||
Db 20 TAGAGGAGAGCAACCAAGTAAGAAAA 49

RESULT 14
A57961 54 bp DNA PAT 05-MAR-1998
LOCUS
DEFINITION Sequence 27 from Patent EP0743364.
ACCESSION A57961
VERSION A57961.1 GI:3713731
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 54)
AUTHORS Narwa, R. and Rogues, P.
TITLE Nucleic acid fragments derived from the HIV-1 genome, corresponding fragments and their application as reactives for risk evaluation of HIV-1 mother-to-fetal transmission
JOURNAL Patent: EP 0743364-A 27 20-NOV-1996;
COMMENT COMMISSARIAT ENERGIE ATOMIQUE (FR)
FEATURES Other publication FR 2734281 961122.
SOURCE 1..54
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 30 a 4 c 15 g 5 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 54;
Best Local Similarity 73.3%; Pred. No. 6.4e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11 tagagcccaagcaatcaacggttaagaca 40
||||| ||||| ||| ||||| |||
Db 20 TAGAGGAGAGCAACCAAGTAAGAAAA 49

RESULT 15
E05823 92 bp DNA PAT 29-SEP-1997
LOCUS
DEFINITION DNA encoding HGP-30 of HIV.
ACCESSION E05823
VERSION E05823.1 GI:2174010
KEYWORDS JP 1993301895-A/1.
SOURCE JP 1993301895-A/1.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 92)
AUTHORS Nagaya, A., Takamura, C. and Kamogawa, K.
TITLE HYBRID ANTIGEN PROTEIN, RECOMBINED VIRUS FOR EXPRESSING THE SAME, AND ITS PRODUCTION
JOURNAL Patent: JP 1993301895-A 1 16-NOV-1993;
COMMENT NIPPON ZEON CO LTD
OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1993301895-A/1
PD 16-NOV-1993

PF 22-APR-1992 JP 1992127980
PI NAGAYA ARSUSHI, TAKAMURA CHIZUKO, KAMOGAWA KOICHI PC
C07K15/04,A61K39/00,A61K39/00,A61K39/275,C12N7/01, PC
C12P21/02//C12N15/39,
PC C12N15/62,C12N15/86,(C12P21/02,C12R1:92);
CC strandedness: Double;
CC topology: Linear;
FEATURES Location/Qualifiers
SOURCE 1..92
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 47 a 10 c 22 g 13 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 5; Length 92;
Best Local Similarity 73.3%; Pred. No. 6.5e+03;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11 tagagcccaagcaatcaacggttaagaca 40
||||| ||||| ||| ||||| |||
Db 55 TAGAGGAGAGCAACCAAGTAAGAAAA 84

Search completed: June 4, 2000, 16:05:44
Job time: 27896 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:21 ; Search time 322.35 Seconds
(Without alignments)
31.046 Million cell updates/sec

Title: US-09-164-714-19

Perfect score: 40
Sequence: 1 cactctgcagtagacgccaagcaatcaacgtaagaaca 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues
Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19.2	48.0	54	T43657	HIV-1 matrix prote
2	18.2	45.5	54	T43658	HIV-1 matrix prote
3	18.2	45.5	54	T43650	HIV-1 matrix prote
4	17.2	43.0	54	T43659	HIV-1 matrix prote
5	17.2	43.0	54	T43642	HIV-1 matrix prote
6	17.2	43.0	54	T43643	HIV-1 matrix prote
7	17.2	43.0	54	T43644	HIV-1 matrix prote
8	17.2	43.0	54	T43645	HIV-1 matrix prote
9	17.2	43.0	54	T43646	HIV-1 matrix prote
10	17.2	43.0	54	T43647	HIV-1 matrix prote
11	17.2	43.0	54	T43649	HIV-1 matrix prote
12	17.2	43.0	54	T43652	HIV-1 matrix prote
13	17.2	43.0	54	T43654	HIV-1 matrix prote
14	17.2	43.0	54	T43655	HIV-1 matrix prote
15	17.2	43.0	92	O53311	Hybrid antigen pro
16	16.8	42.0	54	T43648	HIV-1 matrix prote
17	16.8	42.0	54	T43653	HIV-1 matrix prote
18	16.4	41.0	35	V29846	Human CD4 transmem
19	16.2	40.5	54	T43656	HIV-1 matrix prote
20	16.2	40.5	57	T28274	Primer lux-5'-BamH
21	16	40.0	69	O30897	Primer 312-69. New
22	16	40.0	69	O30892	Primer 312-64. New
23	16	40.0	96	T49233	HVRI region of E2
24	15.8	39.5	97	T19416	Human gene signalu
25	15.6	39.0	46	T79510	Primer Streptag o
26	15.6	39.0	47	X52594	Human genome blall
27	15.6	39.0	47	X52596	Human genome blall
28	15.6	39.0	47	X52546	Human genome blall
29	15.6	39.0	47	X52544	Human genome blall
30	15.6	39.0	54	T43653	Human genome blall
31	15.6	39.0	54	T43651	HIV-1 matrix prote
32	15.2	38.0	38	T43641	HIV-1 matrix prote
33	15.2	38.0	38	N94509	Probe for N-termin
34	15.2	38.0	38	V83523	PCR primer used to

PCR primer used to
Oligonucleotide us
Quadruplex/duplex
Fibroblast growth
Fibroblast growth
Primer 13 for 95 k
PCR primer used to
Staphylococcus aur
DNA sequence encod
Modified type of s
DNA sequence of 1m

ALIGNMENTS

RESULT 1
T43657
ID T43657 standard; DNA: 54 BP.
AC T43657;
DT 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment 4541.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
NS Human immunodeficiency virus type 1.
PN Ep-743364-A2.
PD 20-NOV-1996.
PE 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R, Roques P;
DR WPI: 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3: Page 27: 46pp: French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTC or ATC; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-foetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
CC Sequence 54 BP; 30 A; 4 C; 15 G; 5 T;
SO

Query Match 48.0%; Score 19.2; DB 1; Length 54;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 9 agtagcgcagcaatcaacgtaagaaca 40
||||| | ||||| ||| ||||| |
Db 18 AGTAGAGGAGAGCAAAACAAAGTAGAGAA 49

RESULT 2
T43658
ID T43658 standard; DNA: 54 BP.
AC T43658;
DT 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment 2754.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
NS Human immunodeficiency virus type 1.
PN Ep-743364-A2.
PD 20-NOV-1996.
PE 17-MAY-1996; 401084.

PR 18-MAY-1995: FR-005914.
 PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 PI Narwa R. Roques P;
 DR WPI: 96-507733/51.
 PT Human immunodeficiency virus p17 gene fragments, derived proteins
 PT and antibodies - useful for assessing the risk of maternal
 PT transmission of HIV-1 infection
 PS Claim 3; Page 27; 46pp; French.
 CC This sequence is a specifically claimed example of 21-90 nucleotide
 CC long nucleic acid fragments, derived from the gene encoding part of the
 CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
 CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
 CC Y6 = ATG, TTA, CTC, GTA, CTA, CTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
 CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
 CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
 CC = AAA or AAG, then Y6 is not ATG, Y7 is not GAG or GAA, Y8 is not GAA,
 CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
 CC The new nucleic acid fragments can be used as reagents for determining
 CC and assessing the risk of maternal-fetal transmission of HIV-1, using
 CC standard hybridisation or immuno assays. The presence of such sequences
 CC in maternal blood is strongly correlated with transmission of infection.
 SQ Sequence 54 BP; 30 A; 4 C; 15 G; 5 T;

Query Match 45.5%; Score 18.2; DB 1; Length 54;
 Best Local Similarity 74.2%; Pred. No. 67;
 Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 10 gtagagccaagaatcaacgtaagaaca 40
 ||||| | ||||| ||| ||||| |||
 Db 19 GTAGAGCAAGAGCAAAACAAAGTAAGAAAA 49

RESULT 3
 T43650 standard; DNA; 54 BP.
 AC T43650;
 DE 19-AUG-1997 (first entry)
 DT HIV-1 matrix protein p17 gene fragment HAR.
 KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
 KM detection; maternal transmission; hybridisation assay; immunoassay; ss.
 OS Human immunodeficiency virus type 1.
 FT key
 FT mat_peptide 1..54
 FT /*tag= a
 FT /product= HAR_peptide
 PN EP-743364-A2.
 PD 20-NOV-1996.
 PR 17-MAY-1996; 401084.
 PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 PI Narwa R. Roques P;
 DR WPI: 96-507733/51.
 PT P-PSDB: W06612.
 PT Human immunodeficiency virus p17 gene fragments, derived proteins
 PT and antibodies - useful for assessing the risk of maternal
 PT transmission of HIV-1 infection
 PS Claim 3; Page 25; 46pp; French.
 CC This sequence is a specifically claimed example of 21-90 nucleotide
 CC long nucleic acid fragments, derived from the gene encoding part of the
 CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
 CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
 CC Y6 = ATG, TTA, CTC, GTA, CTA, CTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
 CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
 CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
 CC = AAA or AAG, then Y6 is not ATG, Y7 is not GAG or GAA, Y8 is not GAA,
 CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
 CC The new nucleic acid fragments can be used as reagents for determining
 CC and assessing the risk of maternal-fetal transmission of HIV-1, using
 CC standard hybridisation or immuno assays. The presence of such sequences
 CC in maternal blood is strongly correlated with transmission of infection.
 SQ Sequence 54 BP; 29 A; 5 C; 15 G; 4 T;

Query Match 45.5%; Score 18.2; DB 1; Length 54;
 Best Local Similarity 74.2%; Pred. No. 67;
 Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 10 gtagagccaagaatcaacgtaagaaca 40
 ||||| | ||||| ||| ||||| |||
 Db 19 GTAGAGCAAGAGCAAAACAAAGTAAGAAAA 49

RESULT 4
 T43659 standard; DNA; 54 BP.
 AC T43659;
 DE 19-AUG-1997 (first entry)
 DT HIV-1 matrix protein p17 gene fragment 2826.
 KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
 KM detection; maternal transmission; hybridisation assay; immunoassay; ss.
 OS Human immunodeficiency virus type 1.
 FT EP-743364-A2.
 PD 20-NOV-1996.
 PR 17-MAY-1996; 401084.
 PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 PI Narwa R. Roques P;
 DR WPI: 96-507733/51.
 PT Human immunodeficiency virus p17 gene fragments, derived proteins
 PT and antibodies - useful for assessing the risk of maternal
 PT transmission of HIV-1 infection
 PS Claim 3; Page 27; 46pp; French.
 CC This sequence is a specifically claimed example of 21-90 nucleotide
 CC long nucleic acid fragments, derived from the gene encoding part of the
 CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
 CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
 CC Y6 = ATG, TTA, CTC, GTA, CTA, CTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
 CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
 CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
 CC = AAA or AAG, then Y6 is not ATG, Y7 is not GAG or GAA, Y8 is not GAA,
 CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
 CC The new nucleic acid fragments can be used as reagents for determining
 CC and assessing the risk of maternal-fetal transmission of HIV-1, using
 CC standard hybridisation or immuno assays. The presence of such sequences
 CC in maternal blood is strongly correlated with transmission of infection.
 SQ Sequence 54 BP; 30 A; 4 C; 15 G; 5 T;

Query Match 43.0%; Score 17.2; DB 1; Length 54;
 Best Local Similarity 73.3%; Pred. No. 1.7e+02;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagccaagaatcaacgtaagaaca 40
 |||| | ||||| ||| ||||| |||
 Db 20 TTAGAGCAAGAGCAAAACAAAGTAAGAAAA 49

RESULT 5
 T43642 standard; DNA; 54 BP.
 AC T43642;
 DE 19-AUG-1997 (first entry)
 DT HIV-1 matrix protein p17 gene fragment ARI.
 KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
 KM detection; maternal transmission; hybridisation assay; immunoassay; ss.
 OS Human immunodeficiency virus type 1.
 FT EP-743364-A2.
 PD 20-NOV-1996.
 PR 17-MAY-1996; 401084.
 PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 PI Narwa R. Roques P;
 DR WPI: 96-507733/51.
 PT Human immunodeficiency virus p17 gene fragments, derived proteins
 PT and antibodies - useful for assessing the risk of maternal

	RESULT	6
T43643	ID	T43643 standard; DNA; 54 BP.
AC	AC	T43643;
DT	DT	19-AUG-1997 (first entry)
DE	DE	HIV-1 matrix protein p17 gene fragment B01.
KW	KW	Human immunodeficiency virus; matrix protein p17; prognosis; probe; detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS	OS	Human immunodeficiency virus type 1.
FH	FH	key Location/Qualifiers
FT	FT	mat_peptide
FT	FT	1..54
FT	FT	/*tag= a
PD	PD	/product= B01_peptide
PN	PN	EP-743364-A2.
PR	PR	20-NOV-1996.
PR	PR	17-MAY-1996; 401084.
PR	PR	18-MAY-1995; FR-005914.
PA	PA	(COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI	PI	Narwa R. Roques P.
DR	DR	WPI; 96-507733/51.
DR	DR	P-PSDB; W06610.
PT	PT	Human immunodeficiency virus p17 gene fragments, derived proteins and antibodies - useful for assessing the risk of maternal transmission of HIV-1 infection
PT	PT	Claim 3; Page 23; 46pp; French.
PS	PS	This sequence is a specifically claimed example of 21-90 nucleotide long nucleic acid fragments, derived from the gene encoding part of the p17 HIV-1 matrix protein; the claimed nucleic acids must include at least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA; Y6 = ATA, ATT, CTTG, GTA, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or CAA; Y9 = AAA, GTA, CTA, GAA, GAG, GAC, GAT or ACA; Y10 = CAA, CAG or CAG; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5 = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA, Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC. The new nucleic acid fragments can be used as reagents for determining and assessing the risk of maternal-fetoal transmission of HIV-1, using standard hybridisation or immuno assays. The presence of such sequences in maternal blood is strongly correlated with transmission of infection
Sequence	Sequence	54 BP; 30 A; 4 C; 15 G; 5 T;

Query Match	43.0%;	Score 17.2;	DB 1;	Length 54;
Best Local Similarity	73.3%;	Pred. No. 1.7e+02;		
Matches 22;	Conservative	0;	Mismatches 8;	Indels 0;
0y	11 tagagcgcaagcaatcaacggttaagaaca	40		

CC This sequence is a specifically claimed example of 21-50 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein: y5-y6-y7-y8-y9-y10-y11, where y5 = AAA, AAG or GAA;
CC least the sequence: y5-y6-y7-y8-y9-y10-y11, where y5 = AAA, AAG or GAA;
CC y6 = ATA, TTA, CTG, GTA, CTA, GTG or ATG; y7 = GAG or GAA; y8 = GAA or
CC GAA; y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or ACA; y10 = CAA, CAG or
CC AAG; y11 = AAA, AGT, AAT, CAT, AAC or ACC; provided that when y5
CC = AAA or AAG, then y6 is not ATA, y7 is not GAG or GAA, y8 is not GAA,
CC y9 is not GAG or GAA, y10 is not CAA or CAG and y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fœtal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection
90 Sequence 54 BP; 31 A; 4 C; 14 G; 5 T;

QY 11 tagacgccaagcaaatcaacgtaagaaca 40
||| | ||||| || ||||| |
Db 20 TAGAGGAAGAGCAAAACAAAGTAAAGAAA 45

ID T43645 standard; DNA; 54 BP.

DT 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment PAL.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KM detection; maternal transmission; hybridisation assay; immunosay; ss
CS Human immunodeficiency virus type 1.
PR EP-743364-A2.
PN 20-NOV-1996.
PD 17-MAY-1996; 401084.
PF 18-MAY-1995; FR-005914.
PR (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PA Narwa R., Roques P;
PI WPI: 96-507733/51.
DR Human immunodeficiency virus p17 gene fragments, derived proteins
PT and antibodies - useful for assessing the risk of maternal
PT transmission of HIV-1 infection
PS Claim 3; Page 24; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA.

CC Y6 = ATA, TTA, CTG, GTA, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 31 A; 4 C; 13 G; 6 T;

Query Match 43.0%; Score 17.2; DB 1; Length 54;
Best Local Similarity 73.3%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagcgcaagcaaatcaacgtaagaaca 40
DB 20 TAGAGGAGAGCAAAACAAAGTAAAGAAA 49

RESULT 9

T43646 ID T43646 standard; DNA; 54 BP.

AC T43646;

DT 19-AUG-1997 (first entry)

DE HIV-1 matrix protein p17 gene fragment RYO.

KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;

OS detection; maternal transmission; hybridisation assay; immunoassay; ss.

PN Human immunodeficiency virus type 1.

PP BP-743364-A2.

PD 20-NOV-1996.

PE 17-MAY-1996; 401084.

PF 18-MAY-1995; FR-005914.

PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

PI Narwa R. Roques P.

DR WPI; 96-507733/51.

PT Human immunodeficiency virus p17 gene fragments, derived proteins

PT and antibodies - useful for assessing the risk of maternal

PT transmission of HIV-1 infection

PS Claim 3: Page 24; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, GTA, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 32 A; 5 C; 12 G; 5 T;

Query Match 43.0%; Score 17.2; DB 1; Length 54;
Best Local Similarity 73.3%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagcgcaagcaaatcaacgtaagaaca 40
DB 20 TAGAGGAGAGCAAAACAAAGTAAAGAAA 49

RESULT 10

T43647 ID T43647 standard; DNA; 54 BP.

AC T43647;

DT 19-AUG-1997 (first entry)

DE HIV-1 matrix protein p17 gene fragment FLO.

KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;

KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
PN BP-743364-A2.
PD 20-NOV-1996.
PE 17-MAY-1996; 401084.
PF 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa R. Roques P.
DR WPI; 96-507733/51.
PT Human immunodeficiency virus p17 gene fragments, derived proteins

PT and antibodies - useful for assessing the risk of maternal

PT transmission of HIV-1 infection

PS Claim 3: Page 24; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, GTA, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 30 A; 5 C; 14 G; 5 T;

Query Match 43.0%; Score 17.2; DB 1; Length 54;
Best Local Similarity 73.3%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagcgcaagcaaatcaacgtaagaaca 40
DB 20 TAGAGGAGAGCAAAACAAAGTAAAGAAA 49

RESULT 11

T43649 ID T43649 standard; DNA; 54 BP.

AC T43649;

DT 19-AUG-1997 (first entry)

DE HIV-1 matrix protein p17 gene fragment FAL.

KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;

OS detection; maternal transmission; hybridisation assay; immunoassay; ss.

PN Human immunodeficiency virus type 1.

PP BP-743364-A2.

PD 20-NOV-1996.

PE 17-MAY-1996; 401084.

PF 18-MAY-1995; FR-005914.

PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

PI Narwa R. Roques P.

DR WPI; 96-507733/51.

PT Human immunodeficiency virus p17 gene fragments, derived proteins

PT and antibodies - useful for assessing the risk of maternal

PT transmission of HIV-1 infection

PS Claim 3: Page 25; 46pp; French.

CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, GTA, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAC, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP; 31 A; 4 C; 14 G; 5 T;

Query Match	43.0%;	Score 17.2;	DB 1;	Length 54;
Best Local Similarity	73.3%;	Pred. No. 1.7e+02;		
Matches 22;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;
Oy	11	tagacgccaagcaatcaacgtagaaca	40	
Db	20	TAGAGCAAGAGCAAAACAAAGTAAAGAAA	49	
RESULT 12				
ID	T43652	T43652 standard; DNA; 54 BP.		
AC	T43652;			
DT	19-AUG-1997	(first entry)		
DE	HIV-1 matrix protein p17 gene fragment CHET.			
KW	Human immunodeficiency virus; matrix protein p17; prognosis; probe;			
RV	detection; maternal transmission; hybridisation assay; immunoassay; ss.			
OS	Human immunodeficiency virus type 1.			
PN	EP-743364-A2.			
PD	20-NOV-1996.			
PR	17-MAY-1996; 401084.			
PT	18-MAY-1995; FR-005914.			
PA	(COMS) COMMISSARIAT ENERGIE ATOMIQUE.			
PI	Narwa R, Rogues P;			
PI	WPI: 96-507733/51.			
PT	Human immunodeficiency virus p17 gene fragments, derived proteins			
PT	and antibodies - useful for assessing the risk of maternal			
PT	transmission of HIV-1 infection			
PS	Claim 3; Page 26; 46pp; French.			
CC	This sequence is a specifically claimed example of 21-90 nucleotide			
CC	long nucleic acid fragments, derived from the gene encoding part of the			
CC	p17 HIV-1 matrix protein; the claimed nucleic acids must include at			
CC	least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;			
CC	Y6 = ATA, TTA, CTG, GTA, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or			
CC	AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAG, GAT or AGA; Y10 = CAA, CAG or			
CC	CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5			
CC	= AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,			
CC	Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.			
CC	The new nucleic acid fragments can be used as reagents for determining			
CC	and assessing the risk of maternal-fœtal transmission of HIV-1, using			
CC	standard hybridisation or immuno assays. The presence of such sequences			
CC	in maternal blood is strongly correlated with transmission of infection.			
SO	Sequence 54 BP: 30 A; 4 C; 15 G; 5 T;			
Query Match	43.0%;	Score 17.2;	DB 1;	Length 54;
Best Local Similarity	73.3%;	Pred. No. 1.7e+02;		
Matches 22;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;
Oy	11	tagacgccaagcaatcaacgtagaaca	40	
Db	20	TAGAGCAAGAGCAAAACAAAGTAAAGAAA	49	
RESULT 13				
ID	T43654	T43654 standard; DNA; 54 BP.		
AC	T43654;			
DT	19-AUG-1997	(first entry)		
DE	HIV-1 matrix protein p17 gene fragment SIM.			
KW	Human immunodeficiency virus; matrix protein p17; prognosis; probe;			
RV	detection; maternal transmission; hybridisation assay; immunoassay; ss.			
OS	Human immunodeficiency virus type 1.			
PI	Key	Location/Qualifiers		
PI	mat_peptide	1..54		
PI		/tag= a		
PI		/product= SIM_peptide		
PN	EP-743364-A2.			
PD	20-NOV-1996.			
PR	17-MAY-1996; 401084.			
PT	18-MAY-1995; FR-005914.			
PA	(COMS) COMMISSARIAT ENERGIE ATOMIQUE.			

PI Narwa, R, Roques, P:
DR WPI: 96-507733/51.
DR P-PSDB: W06614.
PR Human immunodeficiency virus p17 gene fragments, derived proteins
PR and antibodies - useful for assessing the risk of maternal
PR transmission of HIV-1 infection
PS Claim 3; Page 26; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTG or ATG; Y7 = GAG or AAG; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP: 30 A; 6 C; 14 G; 4 T;

Query Match 43.0%; Score 17.2; DB 1; Length 54;
Best Local Similarity 73.3%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagacgcacgaacatcaacgttaagaca 40
||||| ||||| ||| ||||| |||
Db 20 TAGAGCAGAGCGCAAAACAAAGTAAAGAAA 49

RESULT 14
T43655
ID T43655 standard; DNA; 54 BP.
AC T43655; 19-AUG-1997 (first entry)
DT 19-AUG-1997 (first entry)
DE HIV-1 matrix protein p17 gene fragment MOE.
KW Human immunodeficiency virus; matrix protein p17; prognosis; probe;
KW detection; maternal transmission; hybridisation assay; immunoassay; ss.
OS Human immunodeficiency virus type 1.
FH Key Location/Qualifiers
FT mat_peptide 1..54
FT /*tag= a
FT /product= MOE_peptide
PN EP-743364-A2.
PD 20-NOV-1996.
PR 17-MAY-1996; 401084.
PR 18-MAY-1995; FR-005914.
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
PI Narwa, R, Roques, P:
DR WPI: 96-507733/51.
DR P-PSDB: W06615.
PR Human immunodeficiency virus p17 gene fragments, derived proteins
PR and antibodies - useful for assessing the risk of maternal
PR transmission of HIV-1 infection
PS Claim 3; Page 26; 46pp; French.
CC This sequence is a specifically claimed example of 21-90 nucleotide
CC long nucleic acid fragments, derived from the gene encoding part of the
CC p17 HIV-1 matrix protein; the claimed nucleic acids must include at
CC least the sequence: Y5-Y6-Y7-Y8-Y9-Y10-Y11, where Y5 = AAA, AAG or GAA;
CC Y6 = ATA, TTA, CTG, CTA, GTG or ATG; Y7 = GAG or GAA; Y8 = GAA or
CC AAG; Y9 = ATA, GTA, CTA, GAA, GAG, GAT or AGA; Y10 = CAA, CAG or
CC CGA; Y11 = AAA, AGT, AAG, AAT, CAT, AAC or ACC; provided that when Y5
CC = AAA or AAG, then Y6 is not ATA, Y7 is not GAG or GAA, Y8 is not GAA,
CC Y9 is not GAG or GAA, Y10 is not CAA or CAG and Y11 is not AAT or AAC.
CC The new nucleic acid fragments can be used as reagents for determining
CC and assessing the risk of maternal-fetal transmission of HIV-1, using
CC standard hybridisation or immuno assays. The presence of such sequences
CC in maternal blood is strongly correlated with transmission of infection.
SQ Sequence 54 BP: 30 A; 6 C; 13 G; 5 T;

Query Match 43.0%; Score 17.2; DB 1; Length 54;
 Best Local Similarity 73.3%; Pred. No. 1.7e+02;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11 tagagccaagcaatcaacggtagaaca 40
 ||||| | ||||| ||| ||||| |
 Db 20 TAGAGGAAGAGCAACCAAGTAAGAAAA 49

RESULT 15

053311
 ID 053311 standard; DNA; 92 BP.
 AC 053311.
 DT 09-JUN-1994 (first entry)
 DE Hybrid antigen protein linker sequence #1.
 KW Linker; recombinant virus; hybrid antigen protein; virus;
 structural protein; peptide epitope; antigen; vaccine; ss.
 OS Synthetic
 PN J05301895-A.
 PD 16-NOV-1993.
 PF 22-APR-1992; 127980.
 PR 22-APR-1992; JP-127980.
 PA (JAPG) NIPPON ZEON KK.
 DR WPI: 93-400397/50.
 PT New hybrid antigen protein and recombinant virus, useful for
 active vaccine - comprises virus structural protein and peptide
 epitope contg. below 300 amino acid(s) useful for antigen of
 component vaccine and expressed by recombinant virus
 PS Disclosure: Page 7; 14pp; Japanese.
 CC The sequences given in 053311-12 are linkers which were used in the
 production of a recombinant virus encoding a hybrid antigen protein.
 CC The hybrid antigen protein has a virus structural protein and a peptide
 epitope and consists of less than 300 amino acids. This hybrid
 CC antigen protein may be useful as the antigen component in a vaccine.
 SQ Sequence 92 BP; 47 A; 10 C; 22 G; 13 T;

Query Match 43.0%; Score 17.2; DB 1; Length 92;
 Best Local Similarity 73.3%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11 tagagccaagcaatcaacggtagaaca 40
 ||||| | ||||| ||| ||||| |
 Db 55 TAGAGGAAGAGCAACCAAGTAAGAAAA 84

Search completed: June 4, 2000, 16:24:21
 Job time: 28917 sec

Gencore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:34 ; Search time 4521.53 Seconds
(without alignments)
35.857 Million cell updates/sec

Title: US-09-164-714-19

Perfect score: 40

Sequence: 1 cactctgcagtacgacccaagcaatacaacgtaagaaca 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*
106: gb_est87:*
107: gb_est88:*
108: gb_est89:*
109: gb_est90:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result

Query

No.	Score	Match Length	DB	ID	Description
1	19	47.5	52	48	A1584790 fb83d08.y
2	18.6	46.5	94	31	AA291480 zC40B04.s
3	17.6	44.0	74	43	A1224006 gx02a05.x
4	17.2	43.0	96	21	D45646 HDMG502842
5	16.6	41.5	43	43	A1198732 qf78c08.x
6	16.6	41.0	67	44	A1286661 ub95q02.r
7	16.4	41.0	93	37	AA173869 nv79b04.s
8	16.2	40.5	93	28	AA066146 mm42d04.r
9	15.8	39.5	89	20	D11978 HDMG512E03
10	15.8	39.5	94	20	D12000 HDMG513A08
11	15.6	39.0	76	61	A1872794 km70f12.x
12	15.6	39.0	79	39	AA868541 ak43e11.s
13	15.6	38.5	48	42	AF027883 AFO27883
14	15.4	38.5	64	37	AA715443 nv53b08.r
15	15.4	38.5	70	49	A1659613 tu06b03.x
16	15.4	38.5	84	83	AF088162 Homo sapi
17	15.4	38.5	89	44	A1252064 qv39f11.x
18	15.4	38.5	89	29	AA152894 mg56f02.r
19	15.4	38.5	99	80	AAW85906 lG1_251.G
20	15.4	38.5	100	22	R76016 Y123F09.r1
21	15.4	38.5	100	37	AA680523 T3670.B10
22	15.4	38.5	100	44	A1252291 qv26f04.x
23	15.2	38.0	78	49	F24150 HSPD10213
24	15.2	38.0	82	82	AL013632 F. rubrip
25	15.2	38.0	100	63	A1954642 wq34e02.x
26	15	37.5	34	43	A1218040 qh29a03.x
27	15	37.5	59	63	A1988246 sc98a02.Y
28	15	37.5	64	27	AA041720 mj05b07.r
29	15	37.5	64	27	W99959 mg29g03.r1
30	15	37.5	67	30	AA218242 mv74d04.r
31	15	37.5	73	48	A1569601 tC44c05.x
32	15	37.5	76	37	AA711088 vE55e05.r
33	15	37.5	78	34	AA512518 vJ18h04.r
34	15	37.5	84	41	A1053592 v18h09.x
35	15	37.5	85	40	A1688982 tX91f08.x
36	15	37.5	100	41	A1061242 an36c05.x
37	14.8	37.0	65	44	A1271266 qn82g05.x
38	14.8	37.0	91	25	N84555 J0022F.Huma
39	14.8	37.0	93	74	AW192657 tu148d03.x
40	14.8	37.0	97	62	A1918993 tu16b08.x
41	14.8	37.0	100	34	AA515148 ng68c05.s
42	14.6	36.5	58	82	HSMC37C10 HSMC10
43	14.6	36.5	63	38	AA746400 nv62c07.s
44	14.6	36.5	63	84	B04181 CSRL-28d10-
45	14.6	36.5	64	27	AA052614 mc82a09.r

RESULT	1
LOCUS	A1584790
DEFINITION	A1584790 52 bp mRNA EST 06-Apr-1999 lB03g08.y_1 zebrafish Washo WPMG EST Danio rerio cDNA 5' similar to SM:FRR-MOUSE Q62356 FOLLISTATIN-RELATED PROTEIN PRECURSOR ; , mRNA sequence.
ACCESSION	A1584790
VERSION	A1584790
KEYWORDS	A1584790.1 GI:4570687
SOURCE	EST. zebrafish.
ORGANISM	Danio rerio Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinoidae; Cyprinidae; Rasbora; Danio.
REFERENCE	1 (bases 1 to 52) Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Bddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,F., Underwood,K., Streptoe,M., Theisling,B., Allen,M., Bowers,Y., Petson,B., Swaller,T., Gibbons,W., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE	Washu zebrafish EST Project 1998					
COMMENT	Unpublished (1998) On May 18, 1998 this sequence version replaced gi:3138646. Contact: Stephen L. Johnson Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: szorist@wustl.edu cDNA Library Preparation: Matthew Clark. cDNA library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and Resourcenetcenter@primardatenbank, Berlin, Germany (web address: www.rzp.de)					
FEATURES	Trace covered overall poor quality Possible reversed clone: similarity on wrong strand Seq primer: T3 ET from Amershams High quality sequence stop: 1.					
SOURCE	location/Dualifiers 1..52 /organism="Danio rerio" /db_xref="taxon:7955" /clone_id="Zebrafish Washu MPMG EST" /sex="mixed" /tissue_type="26 somite embryos, adult livers, shield stage embryos" /lab_host="Xrl-blue MRP" /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer [5'-PGACTAGTCCTAGATCGCAGCCGCCGCCCTTTTATTATTTTATTTT-3']; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT vector (BRL). Library was constructed by Matthew Clark (Lehrbach lab, ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."					
BASE COUNT	15 a 9 c 16 g 12 t					
ORIGIN						
Query Match	47.5%; Score 19; DB 48; Length 52;					
Best Local Similarity	81.5%; Pred. No. 4.2e+02;					
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;						
Oy	7 gcagtagacgcgaagaatcaacggt 33 					
Dd	7 GCAGTAGGCCTCAACGAATCAACGCT 33					
RESULT 2						
AA291480	EST 16-MAY-1997					
LOCUS						
DEFINITION	AA291480 94 bp mRNA					
IMAGE:724783 3'	Soares ovary tumor NBH07 Homo sapiens cDNA clone					
POLYPTIDE II ;	similar to TR:G14016 G14016 CYTOCHROME C OXIDASE					
AA291480						
ACCESSION	AA291480.1 GI:1939475					
VERSION	EST.					
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;						

[illegible]

	unknown library type		Std Error:	0.00
	Insert Length:	638		
	Seq primer:	-40bp from Glibco.		
FEATURES	Location/Qualifiers			
source	1..74			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone_lib="IMAGE:2000144"			
	/clone_lib="NCI-CCAP_Lym12"			
	/tissue_type="lymphoma, follicular mixed small and large cell"			
	/lab_host="DH10B"			
	/note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.25 kb. Life Technologies catalog #: 11547-015"			
BASE COUNT	19 a	17 c	14 g	24 t
ORIGIN				
Query Match	44.0%; Score 17.6;	DB 43;	Length 74;	
Best Local Similarity	71.9%;	Pred. No. 1.7e+03;		
Matches	23; Conservative	0; Mismatches	9; Indels	0; Gaps
OY	8	cagtagacgccgaatcaactaacggttaagaac	39	
Dd	41	CAGTAGACGTCGAAGGGCATTAAAGTCAGAAGAC	10	
RESULT	4			
D45646	D45646	96 bp mRNA EST	20-FEB-1995	
LOCUS	HUMGS02842 Human adult lung 3'	directed MbOI CDNA Homo sapiens cdna		
DEFINITION	3', mRNA sequence.			
ACCESSION	D45646			
VERSION	D45646.1 GI:662600			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 96)			
AUTHORS	Itoh,K., Okubo,K., Yosii,J., Yokouchi,H. and Matsubara,K.			
TITLE	An expression profile of active genes in human lung			
JOURNAL	DNA Research 1, 279-287 (1994)			
MEDLINE	95236275			
COMMENT	Contact: Kohichi Itoh Institute for Molecular and Cellular Biology Osaka University 3-1, Yamadaoka, Suita, Osaka, 565, Japan Tel: 06-877-5111 x3910 Fax: 06-877-1922 Insert Length: 396 Std Error: 0.00 High quality sequence stop: 36. Location/Qualifiers			
FEATURES	1..96			
source	/organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="Human adult lung 3' directed MbOI CDNA" /note="Adult human lung, 3' directed MbOI"			
BASE COUNT	23 a	36 c	14 g	20 t
ORIGIN				
Query Match	43.0%; Score 17.2;	DB 21;	Length 96;	
Best Local Similarity	71.0%;	Pred. No. 2.6e+03;		
Matches	22; Conservative	0; Mismatches	9; Indels	0; Gaps
OY	1	cactctgcagttagaccgccaatcacacg	31	
Dd	63	CTCTCTTCATTAACCACAAGCATCTCAANG	93	

RESULT	5	AI198732/c	AI198732	43 bp	mRNA	EST	27-JAN-1999
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
COMMENT							
FEATURES							
source							
Query Match							
Best Local Similarity							
Matches							
QY							
Db							
RESULT							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							

REFERENCE	1 (bases 1 to 67)									
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucata,T., Lacy,M., Le,M., Martin,T., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.									
TITLE	The WashU-HHMI Mouse EST Project									
JOURNAL	Unpublished (1996)									
COMMENT	On Dec 5, 1997 this sequence version replaced g1:2662749. Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicinep 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.edu This clone is available royalty-free through LINT ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MG1:907990									
FEATURES	Trace considered overall poor quality Possible reversed clone: similarity on wrong strand Seq primer: -28ml3 rev2 ET from Amersham High quality sequence stop: 1.									
SOURCE	Location/Qualifiers									
	1..67									
	/organism="Mus musculus"									
	/strain="C57BL/6J"									
	/db_xref="taxon:10090"									
	/clone="IMAGE:1396274"									
	/clone_lib="Soares_mammary_gland_Nbwmg"									
	/sex="male"									
	/tissue_type="mammary gland"									
	/dev_stage="4 weeks"									
	/lab_host="DH10B"									
	/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15' TGTTACCAATCTGATGAGGAGCGCCGCGAATGGTTTTTTTTTTTTTTTTTT T 3'] : double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. library constructed and normalized by Bento Soares and M.Fatima Bernaldo."									
BASE COUNT	12 a	14 c	18 g	23 t						
ORIGIN										
Query Match	41.5%	Score 16.6;	DB 44;	Length 67;						
Best Local Similarity	71.0%	Pred. No. 4.3e+03;								
Matches 22; Conservative	0;	Mismatches 9;	Indels 0;	Gaps 0;						
OY	10	gtagacgcccaagcaatcaacgtaagaca	40							
Db	46	GAGAGCGCCATCCGATCAACGTCACATCA	16							
RESULT 7										
LOCUS	AA713869/c									
DEFINITION	AA713869	93 bp	mrna	EST	29-DEC-1997					
	nv79BD04.s1	NCI_CGAP_Br4	Homo sapiens	cDNA clone IMAGE:1235983						
	similar to	gb:X57809	IG LAMBDA CHAIN C REGIONS (HUMAN);	mrna						
ACCESSION	AA713869									
VERSION	AA713869.1	GI:2726143								
KEYWORDS	EST.									
SOURCE	human.									
ORGANISM	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;									
	Eutheria; Primates; Catarrhini; Homnidae; Homo.									
REFERENCE	1 (bases 1 to 93)									
AUTHORS	NCI-CGAP	ntcpc://www.ncbi.nlm.nih.gov/ncicgap.								

JOURNAL	COMMENT	FEATURES	BASE COUNT	ORIGIN
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	Unpublished (1997)			
On Sep 12, 1996 this sequence version replaced gi:1407147.	Contact: Robert Strausberg, Ph.D.			
Tel: (301) 496-1550	Email: Robert.Strausberg@nih.gov			
Tissue Procurement: Ilan Kirsch, M.D., Kristina A. Cole, M.D., Ph.D. student, Rodrigo F. Chuagui, M.D., Michael R. Emmert-Buck, M.D., Ph.D.	CDNA Library Preparation: David B. Krizman, Ph.D.			
CDNA Library Arrayed by: Greg Lennon, Ph.D.	DNA Sequencing by: Washington University Genome Sequencing Center			
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/URL at:	www.bio.llnl.gov/bdrrp/image/image.html			
Trace considered overall poor quality	Seq primer: -40ml3 fwd, RT from Amersham			
High quality sequence stop: 1.	Location/Qualifiers			
1..93	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:1235983"			
	/clone_1lb="NCI-CGAP_Br4"			
	/sex="female"			
	/tissue_type="normal ductal tissue"			
	/lab_host="DH10B"			
	/note="Organ: breast; Vector: pAMP10; mRNA made from normal breast ductal tissue; cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."			
8 a	25 c	27 g	33 t	
Query Match	41.0%; Score 16.4; DB 37; Length 93;			
Best Local Similarity	67.6%; Pred. No. 5.5e+03;			
Matches	23; Conservative 0; Mismatches 11; Indels 0; Gaps 0;			
QY	7 gcagtagagccgaagcaatacaagcgtgtagaca 40			
Db	35 GCAGTGAGACCAACACACACCTCAAAAACA 2			
RESULT 8				
AA066146/c	93 bp mRNA EST			
LOCUS	mm42d04.11 StrataGene mouse melanoma (#937312) Mus musculus cDNA			
DEFINITION	clone IMAGE:524167 5' similar to gb:U0049 Mouse mRNA for mouse			
ACCESSION	melanoma antigen, complete cds (MOUSE);, mRNA sequence.			
VERSION	AA066146			
KEYWORDS	AA066146.1 GI:1563134			
ORGANISM	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
AUTHORS	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
	1 (bases 1 to 93)			
	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,			
	Geisel,S., Kucada,T., Lacy,M., Le,M., Martin,J., Morris,M.,			
	Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,			
	Thising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and			
	Waterston,R.			
	The WashU-HMT Mouse EST Project			
	Unpublished (1996)			
	On Apr 14, 1993 this sequence version replaced gi:692360.			
	Contact: Marra M/Mouse EST Project			
	WashU-HMT Mouse EST Project			
	Washington University School of Medicine			
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108			

```

Tel: 314 286 1800
Fax: 314 286 1810
Email: mouse@est.watson.wustl.edu
This clone is available royalty-free through LINT; contact the
IMAGE Consortium (infoimage.lint.gov) for further information.
MGI:318015
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
    1..93
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        /clone="IMAGE:524167"
        /clone_lib="Stratagene mouse melanoma (#937312)"
        /tissue_type="melanoma"
        /dev_stage="M2 cells"
        /lab_host="SOLR (kanamycin resistant)"
        /note="Organ: skin; Vector: pBluescript SK-; Site_1:
        EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
        Oligo dT. From M2 cells, a highly metastatic derivative of
        the K-1735 (mouse) melanom. Average insert size: 1.0 kb."
        Uni_ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCCACGAG
        3' -3' adaptor sequence: 5' CTCGACTTTTCTTTTTTTTTT 3'"
BASE COUNT      32 a          16 c          22 g          23 t
ORIGIN
Query Match      40.5%, Score 16.2; DB 28; Length 93;
Best Local Similarity 72.4%; Pred. No. 6.7e+03;
Matches 21; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Oy      5      ctgcagtagacgccaagaatcaacagt 33
         |||||
Db       86      CTCGACTTGCTCTTAGCTTATGAAGGT 58
RESULT 9
D11978/c
LOCUS      D11978      89 bp      mRNA      EST      02-DEC-1992
DEFINITION HUM0512E03 Liver HepG2 cell line. Homo sapiens cDNA clone s12e03,
            mRNA sequence.
ACCESSION   D11978
VERSION     D11978.1 GI:2148809
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 89)
Okubo,K., Hori,N., Matoba,R., Niiyama,T., Fukushima,A., Kojima,Y.,
and Matsubara,K.
Large scale cDNA sequencing for analysis of quantitative and
qualitative aspects of gene expression
Nature Genet. 2, 173-179 (1992)
94258199
JOURNAL     Contact: Kensaku Okubo, Naohito Hori, Ryo Matoba, Toshiyuki
MEDLINE     Niijama, Atsushi Fukushima, Yuko Kojima & Kenichi Matsubara
COMMENT      Institute for Molecular and Cellular Biology
Osaka University
1-3 Yamada-Oka,Suite,Osaka 565,Japan.
FEATURES
source
    1..89
        /organism="Homo sapiens"
        /db_xref="GDB:D058180E"
        /db_xref="taxon:9606"
        /clone="s12e03"
        /clone_lib="Liver HepG2 cell line."
        /lab_host="E.coli"
        /note="3'-directed regional cDNA library. Cleaved by MboI
        and transformed into E.coli."
BASE COUNT      14 a          21 c          22 g          32 t
ORIGIN

```

QY	6	tgcagtagacgccaagcaatcaacgtgaagaaca	40
Db	79	TACCGGAATGCGCAGCAATCATCACAGGAGAGACA	45
RESULT	10		
D12000/c			
LOCUS	D12000	94 bp	mRNA EST 02-DEC-1992
DEFINITION	HUM0S13A08 Liver HepG2 cell line. Homo sapiens cDNA clone sl3a08,		
ACCESSION	D12000		
VERSION	D12000.1	GI:2148823	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
AUTHORS	Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	Okubo, K., Horii, N., Matoba, R., Niijama, T., Fukushima, A., Kojima, Y. and Matsubara, K. Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression Nature Genet. 2, 173-179 (1992)		
JOURNAL	Contact: Kousaku Okubo, Naohiro Horii, Ryo Matoba, Toshiyuki Niijama, Atsushi Fukushima, Yuko Kojima & Kenichi Matsubara		
MEDLINE	Institute for Molecular and Cellular Biology		
COMMENT	Osaka University 1-3 Yamada-oka, Suita, Osaka 565, Japan. Location/Qualifiers 1..94 /organism="Homo sapiens" /db_xref="GDB:D058194E" /db_xref="taxon:9606" /clone="sl3a08" /clone_1lb="Liver HepG2 cell line." /lab_host="E.coli" /note="3'-directed regional cDNA library. Cleaved by MboI and transformed into E.coli."		
BASE COUNT	18 a	21 c	21 g 33 t 1 others
ORIGIN			
Query Match		39.5%; Score 15.8; DB 20; Length 94;	
Best Local Similarity		65.7%; Pred. No. 9.7e+03;	
Matches	23; Conservative	0; Mismatches 12; Indels	0; Gaps 0;
QY	6	tgcagtagacgccaagcaatcaacgtgaagaaca	40
Db	79	TACCGGAATGCGCAGCAATCATCACAGGAGAGACA	45
RESULT	11		
D12000/c			
LOCUS	A1872794	76 bp	mRNA EST 30-AUG-1999
DEFINITION	wm70f12.x1 NCI-CGAP.Ut2 Homo sapiens cDNA IMAGE:2441327 3'		
ACCESSION	A1872794		
VERSION	A1872794.1	GI:5546843	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
AUTHORS	Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		

JOURNAL	Unpublished (1997)
COMMENT	On Dec 20, 1995 this sequence version replaced g1:1134490. Contact: Robert Strausberg, Ph.D. Tel.: (301) 496-1550 Email: Robert_Strausberg@nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmer-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Greg Lennon, Ph.D. Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.lnl.gov/bdrr/image/image.html
FEATURES	Seq primer: -40UP from GIBCO High quality sequence stop: 30.
SOURCE	Location/Qualifiers 1..76 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2441327" /clone_lib="NCI CGAP Jnt2" /tissue_type="moderately-differentiated endometrial adenocarcinoma, 3 pooled tumors" /lab_host="DH10B" /note="Organ: uterus; Vector: PCMV-SPOrt6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.85 kb. Life Technologies catalog #: 11539-012"
BASE COUNT	11 a 22 c 11 g 32 t
ORIGIN	
Query Match	39.0%; Score 15.6; DB 61; Length 76;
Best Local Similarity	63.2%; Pred. No. 1,1e+04;
Matches 24; Conservative 0; Mismatches 14; Indels 0; Gaps 0;	
Db 51 AATTTGAAGTCGAAGAAGAAAGAACCCAAAGAAC 14	2 actctgcagtagacgccaacatcaactgaagaac 39
RESULT 12	
LOCUS	AAB68541 79 bp mRNA EST 04-JAN-1999
DEFINITION	a4k3e11.s1 Soares_Teslis_NHT Homo sapiens CDNA Clone IMAGE:1408748
ACCESSION	AAB68541
VERSION	AAB68541.1 GI:2963986
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 79) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)
JOURNAL	On Jan 19, 1998 this sequence version replaced g1:2152084. Contact: Robert Strausberg, Ph.D. Tel.: (301) 496-1550 Email: Robert_Strausberg@nih.gov CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Boraldo, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.lnl.gov/bdrr/image/image.html
COMMENT	Possible reversed clone: similarity on wrong strand insert length: 1180 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1.79

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1408748"
/clone_1ib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5', TGTTCACATCTGAAGTGGAGCGGCCCAATTTTGTGTGTGT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 8 a 21 c 24 g 26 t

ORIGIN

Query Match 39.0%; Score 15.6; DB 39; Length 79;
Best Local Similarity 70.0%; Pred. No. 1.2e+04;
Matches 21; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 11 tagacgcaagcaatcaacgtaagaaca 40
| | | | | | | | | | | | | | | | | |
Db 78 TGCTCCGCAAGCAAGATCCGTATGCCA 49

RESULT 13
AF027883 84 bp mRNA EST 29-SEP-1998
LOCUS AF027883 Mouse cDNA (S.Herblot) Mus musculus cDNA clone
DEFINITION 18,111..2,104, mRNA sequence.
ACCESSION AF027883
VERSION AF027883.1 GI:3641183
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Herblot, S., Chastagner, P., Samady, L., Moreau, J.-L., Froussard, P.,
Bonnet, J. and Thèze, J. 1 (bases 1 to 84)

TITLE mRNA level analysis of IL-2-induced genes selected by cDNA
Unpublished (1998)
JOURNAL On Jan 19, 1998 this sequence version replaced gi:2153498.
COMMENT Contact: Herblot, Sabine
Immunogenetique Cellulaire
Institut Pasteur
28 rue du Dr. Roux, Paris 75 724, France
Email: sabine.herblot@immol.u-bordeaux2.fr
IL-2-induced mRNA.

FEATURES Location/Qualifiers

source 1..84
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="18,111..2,104"
/clone_1ib="Mouse cDNA (S.Herblot)"
BASE COUNT 11 a 21 c 16 g 36 t

ORIGIN

Query Match 39.0%; Score 15.6; DB 42; Length 84;
Best Local Similarity 70.0%; Pred. No. 1.2e+04;
Matches 21; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 cactctgcaatagacgccaagcaatcaac 30
| | | | | | | | | | | | | | | | | |
Db 84 CACGACGACAGACAGACGAGTAAACAC 55

RESULT 14
AA715443 64 bp mRNA EST 22-JAN-1998
LOCUS AA715443/3 n53b08.r1 NCI-CGAP_Ew1 Homo sapiens cDNA clone IMAGE:1233495
DEFINITION Similar to TR:000193 000193 SMALL ACIDIC PROTEIN. ;, mRNA sequence.
ACCESSION AA715443
VERSION AA715443.1 GI:2727717
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 64)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbdp/image/image.html

Trace considered overall poor quality
Insert Length: 465 Std Error: 0.00
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES

source

1..64
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1233495"
/clone_1ib="NCI-CGAP_Ew1"
/tissue_type="Ewing's sarcoma"
/lab_host="DH10B"
/note="Vector: PAMP10; mRNA made from Ewing's sarcoma,
cDNA made by oligo-dT priming. Non-directionally cloned.
Size-selected on agarose gel, average insert size 600 bp.
Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."

BASE COUNT 10 a 14 c 16 g 24 t

ORIGIN

Query Match 38.5%; Score 15.4; DB 37; Length 64;
Best Local Similarity 76.0%; Pred. No. 1.3e+04;
Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 10 gttagcgcaagcaatcaacgta 34
| | | | | | | | | | | | | | | | | |
Db 62 GAAGACGACAGAGAAATCTACCGAA 38

RESULT 15
A1659613 70 bp mRNA EST 10-MAY-1999
LOCUS A1659613/3 t06b03.x1 NCI-CGAP_Pt28 Homo sapiens cDNA clone IMAGE:2250221 3'
DEFINITION Similar to TR:075918 075918 4F5REL. ;, mRNA sequence.
ACCESSION A1659613
VERSION A1659613.1 GI:4763183
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 70)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)
On Mar 16, 1998 this sequence version replaced gi:2961802.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

FEATURES

source

Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers

```
1..70
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2250221"
/clone_lib="NCI-CCAP_Pr28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: p77T3D-Pac (Pharmacia)
with a modified polylinker. Plasmid DNA from the
normalized library NCI-CCAP_Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonesids
985608-986759, 1101192-1101959, and 1217928-1220615)."
Subtraction by Bento Soares and M. Fatima Bonaldo.
```

BASE COUNT 3 a 23 c 14 g 30 t
ORIGIN

Query Match

Best Local Similarity 66.7%; Score 15.4; DB 49; Length 70;
Matches 22; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 6 tagcagtagacgccaagcaatcaacggtagaag 38
||||| ||| || | || ||||| ||||
Db 42 TGCACGACAGACGAGAAAAAGCGACGACGAAAAA 10

Search completed: June 4, 2000, 13:53:37
Job time: 20961 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:48 ; Search time 244.64 Seconds
(without alignments)
21.253 Million cell updates/sec

Title: US-09-164-714-19

Perfect score: 40

Sequence: 1 cactctgcagtagacgcgaagcaatcaacgctagaaca 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 373880

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/6.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/PCITUS.COMB.seq:*
7: /cgn2_6/ptodata/2/1na/Backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19.2	48.0	54	US-08-649-991-25	Sequence 25, Appl
2	18.2	45.5	54	US-08-649-991-18	Sequence 18, Appl
3	18.2	45.5	54	US-08-649-991-26	Sequence 26, Appl
4	18.2	45.5	54	US-08-649-991-120	Sequence 120, Appl
5	17.2	43.0	54	US-08-649-991-10	Sequence 10, Appl
6	17.2	43.0	54	US-08-649-991-11	Sequence 11, Appl
7	17.2	43.0	54	US-08-649-991-12	Sequence 12, Appl
8	17.2	43.0	54	US-08-649-991-13	Sequence 13, Appl
9	17.2	43.0	54	US-08-649-991-14	Sequence 14, Appl
10	17.2	43.0	54	US-08-649-991-15	Sequence 15, Appl
11	17.2	43.0	54	US-08-649-991-17	Sequence 17, Appl
12	17.2	43.0	54	US-08-649-991-20	Sequence 20, Appl
13	17.2	43.0	54	US-08-649-991-22	Sequence 22, Appl
14	17.2	43.0	54	US-08-649-991-23	Sequence 23, Appl
15	17.2	43.0	54	US-08-649-991-27	Sequence 27, Appl
16	17.2	43.0	54	US-08-649-991-119	Sequence 119, Appl
17	17.2	43.0	54	US-08-649-991-121	Sequence 121, Appl
18	17.2	43.0	54	US-08-649-991-122	Sequence 122, Appl
19	17.2	43.0	54	US-08-649-991-123	Sequence 123, Appl
20	17.2	43.0	54	US-08-649-991-124	Sequence 124, Appl
21	17.2	43.0	54	US-08-649-991-125	Sequence 125, Appl
22	17.2	43.0	54	US-08-649-991-126	Sequence 126, Appl
23	17.2	43.0	54	US-08-649-991-128	Sequence 128, Appl
24	17.2	43.0	54	US-08-649-991-129	Sequence 129, Appl
25	16.8	42.0	54	US-08-213-811-10	Sequence 10, Appl
26	16.8	42.0	54	US-08-649-991-16	Sequence 16, Appl
27	16.8	42.0	54	US-08-649-991-21	Sequence 21, Appl

28	16.2	40.5	54	US-08-649-991-24	Sequence 24, Appl
29	16.2	40.5	57	US-08-316-950-9	Sequence 9, Appl
30	16.2	40.5	57	PCT-US95-12642-9	Sequence 9, Appl
31	16	40.0	69	US-07-916-098A-26	Sequence 26, Appl
32	16	40.0	69	US-07-916-098A-26	Sequence 31, Appl
33	15.8	39.5	35	US-08-213-811-9	Sequence 9, Appl
34	15.8	39.5	57	US-08-213-811-9	Sequence 9, Appl
35	15.8	39.5	7	5514566-17	Patent No. 5514566
36	15.6	39.0	54	US-08-649-991-23	Sequence 23, Appl
37	15.6	39.0	54	US-08-649-991-19	Sequence 19, Appl
38	15.6	39.0	54	US-08-649-991-31	Sequence 31, Appl
39	15.4	38.5	65	US-08-213-811-6	Sequence 6, Appl
40	15	37.5	35	US-08-811-492-93	Sequence 93, Appl
41	15	37.5	60	PCT-US95-11985A-11	Sequence 11, Appl
42	15	37.5	69	US-08-096-623A-42	Sequence 42, Appl
43	15	37.5	71	US-08-095-726-36	Sequence 36, Appl
44	15	37.5	71	US-08-096-043-33	Sequence 33, Appl
45	15	37.5	71	US-08-093-577-29	Sequence 29, Appl

ALIGNMENTS

```
RESULT 1
; Sequence 25, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-649-991-25
Query Match 48.0%; Score 19.2; DB 4; Length 54;
Best Local Similarity 75.0%; Pred. No. 13;
```

Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 9 agtagacgccaagcaaatcaacgtaagaaca 40
||||| | ||||| ||| ||||| |
Db 18 AGTAGAGCAAGCAAAACAAAGTAAGAAAA 49

RESULT 2

US-08-649-991-18
; Sequence 18, Application US/08649991
; Patent No. 5919462

; GENERAL INFORMATION:

; APPLICANT: Narva, Remy

; APPLICANT: Roques, Pierre

; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE

; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR

; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF

; NUMBER OF SEQUENCES: 130

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP

; STREET: 1800 M Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20036-5869

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/649,991

; FILING DATE: 17-MAY-1996

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: FR 9505914

; FILING DATE: 18-MAY-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Adler, Reid G.

; REGISTRATION NUMBER: 30,988

; REFERENCE/DOCKET NUMBER: ORES-5003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-467-7000

; TELEFAX: 202-467-7176

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 54 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-649-991-18

Query Match 45.5%; Score 18.2; DB 4; Length 54;
Best Local Similarity 74.2%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 10 gtagacgccaagcaaatcaacgtaagaaca 40
||||| | ||||| ||| ||||| |
Db 19 GTAGAGCAAGCAAAACAAAGTAAGAAAA 49

RESULT 3

US-08-649-991-26

; Sequence 26, Application US/08649991

; Patent No. 5919462

; GENERAL INFORMATION:

; APPLICANT: Narva, Remy

; APPLICANT: Roques, Pierre

; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE

; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR

; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF

; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1

; NUMBER OF SEQUENCES: 130

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP

; STREET: 1800 M Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20036-5869

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/649,991

; FILING DATE: 17-MAY-1996

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: FR 9505914

; FILING DATE: 18-MAY-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Adler, Reid G.

; REGISTRATION NUMBER: 30,988

; REFERENCE/DOCKET NUMBER: ORES-5003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-467-7000

; TELEFAX: 202-467-7176

; INFORMATION FOR SEQ ID NO: 26:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 54 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-649-991-26

Qy 10 gtagacgccaagcaaatcaacgtaagaaca 40
||||| | ||||| ||| ||||| |
Db 19 GTAGAGCAAGCAAAACAAAGTAAGAAAA 49

RESULT 4

US-08-649-991-120

; Sequence 120, Application US/08649991

; Patent No. 5919462

; GENERAL INFORMATION:

; APPLICANT: Narva, Remy

; APPLICANT: Roques, Pierre

; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE

; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR

; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF

; NUMBER OF SEQUENCES: 130

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP

; STREET: 1800 M Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20036-5869

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 120:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-649-991-120

Query Match          45.5%; Score 18.2; DB 4; Length 54;
Best Local Similarity 74.2%; Pred. No. 33;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 10 gtacgcgaagcaatcaacgtaagaca 40
      ||||| ||||| ||| ||||| |
Db 19 gtacagagagcaaaacaaagtaagaaa 49

RESULT 5
US-08-649-991-10
; Sequence 10, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; City: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-649-991-10

Query Match          43.0%; Score 17.2; DB 4; Length 54;
Best Local Similarity 73.3%; Pred. No. 85;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11 tagacgcgaagcaatcaacgtaagaca 40
      ||||| ||||| ||| ||||| |
Db 20 tagacagagagcaaaacaaagtaagaaa 49

RESULT 6
US-08-649-991-11
; Sequence 11, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; City: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-649-991-11

Query Match          43.0%; Score 17.2; DB 4; Length 54;
Best Local Similarity 73.3%; Pred. No. 85;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

RESULT 7
US-08-649-991-12
; Sequence 12, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-649-991-12

Query Match 43.0%; Score 17.2; DB 4; Length 54;
Best Local Similarity 73.3%; Pred. No. 85;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 11 tagagcgcaagcaatcaacggttaagaaca 40
||||| ||||| ||| ||||| |
Db 20 TAGAGCAAGACCAAAACAAAGTAAGAAAA 49

RESULT 8
US-08-649-991-13
; Sequence 13, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington

STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-13

Query Match 43.0%; Score 17.2; DB 4; Length 54;
Best Local Similarity 73.3%; Pred. No. 85;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 11 tagagcgcaagcaatcaacggttaagaaca 40
||||| ||||| ||| ||||| |
Db 20 TAGAGCAAGACCAAAACAAAGTAAGAAAA 49

RESULT 9
US-08-649-991-14
; Sequence 14, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:

NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-14

Query Match 43.0%; Score 17.2; DB 4; Length 54;
Best Local Similarity 73.3%; Pred. No. 85;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagcgaagcaaatcaacgtaagaca 40
||||| ||||| ||| ||||| ||
Db 20 TAGAGGAAGACAAACAAAGTAAGAAA 49

RESULT 10
US-08-649-991-15
Sequence 15, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Narwa, Remy
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-15

Query Match 43.0%; Score 17.2; DB 4; Length 54;
Best Local Similarity 73.3%; Pred. No. 85;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagcgaagcaaatcaacgtaagaca 40
||||| ||||| ||| ||||| ||
Db 20 TAGAGGAAGACAAACAAAGTAAGAAA 49

RESULT 11
US-08-649-991-17
Sequence 17, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Narwa, Remy
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: ORES-5003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-649-991-17

Query Match 43.0%; Score 17.2; DB 4; Length 54;
Best Local Similarity 73.3%; Pred. No. 85;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 11 tagagcgaagcaaatcaacgtaagaca 40
||||| ||||| ||| ||||| ||
Db 20 TAGAGGAAGACAAACAAAGTAAGAAA 49

RESULT 12
US-08-649-991-20
Sequence 20, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Narwa, Remy
APPLICANT: Roques, Pierre

;; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
;; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
;; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
;; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
;; NUMBER OF SEQUENCES: 130
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
;; STREET: 1800 M Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20036-5869
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/649,991
;; FILING DATE: 17-MAY-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: FR 9505914
;; FILING DATE: 18-MAY-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Adler, Reid G.
;; REGISTRATION NUMBER: 30,988
;; REFERENCE/DOCKET NUMBER: ORES-5003
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-467-7000
;; TELEFAX: 202-467-7176
;; INFORMATION FOR SEQ ID NO: 20:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 54 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; US-08-649-991-20

Query Match 43.0%; Score 17.2; DB 4; Length 54;
Best Local Similarity 73.3%; Pred. No. 85;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11 tagagccaaagcaatcaacggttaagaaca 40
||||| ||||| ||| ||||| |
DB 20 TAGAGGAAGACCAAAACAAAGTAAGAAAA 49

RESULT 13
US-08-649-991-22
; Sequence 22, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/649,991
;; FILING DATE: 17-MAY-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: FR 9505914
;; FILING DATE: 18-MAY-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Adler, Reid G.
;; REGISTRATION NUMBER: 30,988
;; REFERENCE/DOCKET NUMBER: ORES-5003
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-467-7000
;; TELEFAX: 202-467-7176
;; INFORMATION FOR SEQ ID NO: 22:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 54 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; US-08-649-991-22

Query Match 43.0%; Score 17.2; DB 4; Length 54;
Best Local Similarity 73.3%; Pred. No. 85;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11 tagagccaaagcaatcaacggttaagaaca 40
||||| ||||| ||| ||||| |
DB 20 TAGAGGAAGACCAAAACAAAGTAAGAAAA 49

RESULT 14
US-08-649-991-23
; Sequence 23, Application US/08649991
; Patent No. 5919462
; GENERAL INFORMATION:
; APPLICANT: Narwa, Remy
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
; TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
; TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
; TITLE OF INVENTION: MATERNOFETAL TRANSMISSION OF HIV-1
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,991
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9505914
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: ORES-5003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 23:

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:05:44 ; Search time 1236.38 Seconds

(without alignments)
-31.472 Million cell updates/sec

Title: US-09-164-714-20

Sequence: 1 gcatgtcgcgttagatgagctacaagcgctgatttgat 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 segs, -486395729 residues

Total number of hits satisfying chosen parameters: 356616

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_com:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pl2:*
11: gb_pl3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_hlg1:*
33: gb_hlg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pir4:*
41: gb_hlg3:*
42: gb_hlg4:*
43: gb_hlg5:*
44: gb_hlg6:*

45: gb_hlg7:*
46: em_hlg1:*
47: em_hlg2:*
48: em_hlg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pl5:*
52: gb_hlg6:*
53: gb_hlg9:*
54: gb_hlg10:*
55: gb_hlg11:*
56: gb_hlg12:*
57: gb_hlg13:*
58: gb_hlg14:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.4	43.5	49	5 A00132	A00132 Nucleotide
2	17.4	43.5	54	3 B7BC12YD	X56119 B.Taurus BC
3	16.6	41.5	25	5 A62916	A62916 Sequence 15
4	16	40.0	60	5 I11638	I11638 Sequence 12
5	16	40.0	80	5 A10881	A10881 Synthetic D
6	16	40.0	80	5 A10882	A10882 Synthetic D
7	15.6	39.0	75	9 D50777	D50777 Human mRNA
8	15.2	38.0	35	5 I28730	I28730 Sequence 34
9	15.2	38.0	35	5 I89384	I89384 Sequence 34
10	15.2	38.0	87	34 AGX487	272092 A.gambiae s
11	14.8	37.0	44	5 A85333	A85333 Sequence 13
12	14.6	36.5	22	5 I59462	I59462 Sequence 11
13	14.6	36.5	33	5 AR014522	AR014522 Sequence
14	14.6	36.5	60	14 SYNANVAA	M60029 Avian neov1
15	14.6	36.5	62	1 J3SPRSEC	M34566 Plasmid pJS
16	14.6	36.5	100	16 RSVRNAL	X67587 Rous sarcom
17	14.4	36.0	45	5 I89301	I89301 Sequence 15
18	14.4	36.0	63	14 SYNMPRT	M12476 Synthetic P
19	14.4	36.0	70	5 A06245	A06245 Synthetic H
20	14.4	36.0	99	34 DROHIG04	D13878 Drosophila
21	14.2	35.5	46	12 M0SIGDJA	L36723 Mus musculu
22	14.2	35.5	57	5 I07481	I07481 Sequence 14
23	14.2	35.5	60	12 M0SIGHTB	M59872 Mouse Ig sw
24	14.2	35.5	73	34 SUPHISAAC	M25332 Sea urchin
25	14.2	35.5	84	5 AR007052	AR007052 Sequence
26	14.2	35.5	84	5 I74992	I74992 Sequence 19
27	14.2	35.5	84	40 S40179	S40179 pancreatic
28	14.2	35.5	99	5 AR022142	AR022142 Sequence
29	14	35.0	33	5 I14299	I14299 Sequence 7
30	14	35.0	35	5 I86947	I86947 Sequence 36
31	14	35.0	55	5 A87313	A87313 Sequence 38
32	14	35.0	57	5 AR015972	AR015972 Sequence
33	14	35.0	57	5 AR050901	AR050901 Sequence
34	14	35.0	57	5 I33401	I33401 Sequence 1
35	14	35.0	67	12 RNTCRAJ32	Y09184 R.norvegicu
36	14	35.0	72	16 REOSAB	J02333 Reovirus se
37	14	35.0	77	6 PT4TRP	K00354 Bacterioph
38	14	35.0	78	4 CIIURSA	K03165 Duck USA sm
39	14	35.0	79	5 AR062563	AR062563 Sequence
40	14	35.0	79	5 I47287	I47287 Sequence 21
41	14	35.0	89	11 H5VAVPG07	AF030207 Homo sap1
42	13.8	34.5	26	5 AR020758	AR020758 Sequence
43	13.8	34.5	26	5 I34039	I34039 Sequence 13
44	13.8	34.5	39	5 A37820	A37820 Sequence 3
45	13.8	34.5	39	5 AR019386	AR019386 Sequence

ALIGNMENTS

RESULT 1
A00132 49 bp DNA PAT 28-JAN-1993
LOCUS Nucleotide sequence 1 from patent number EP0235046.
ACCESSION A00132
VERSION A00132.1 GI:14415
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 49)
AUTHORS Bishop, C., Cotinot, C., Fellous, M., Kirszenbaum, M. and Vaiman, M.
TITLE DNA probe specific for the male genomes of ruminants, their preparation and use
JOURNAL Patent: EP 0235046-A 1 02-SEP-1987;
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE, Etablissement public d'it.; INSTITUT PASTEUR; COMMISSARIAT A L'ENERGIE ATOMIQUE
FEATURES
source 1. .49
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 10 a 10 c 16 g 13 t
ORIGIN

Query Match 43.5%; Score 17.4; DB 5; Length 49;
Best Local Similarity 77.8%; Pred. No. 3.8e+03;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 8 gacgtagatgagctacaagcgctgatt 34
||| ||||| ||| ||||| ||| |||
Db 12 GACCGAGATGTGCTCCACGAGAGTGT 38

RESULT 2
BTBC12YD/c 54 bp DNA MAM 07-APR-1992
LOCUS B.Taurus BCL1.2 Y-specific DNA.
DEFINITION X56119
ACCESSION X56119.1 GI:119
VERSION BCL1.2 sequence; Y-specific sequence.
KEYWORDS Bos taurus.
SOURCE
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 54)
AUTHORS Cotinot, C. and KIRSZENBAUM, M.
TITLE Direct Submission
JOURNAL Submitted (22-SEP-1990) Citinot C., Lab de Biologie Cellulaire et Molculaire, INRA-CRJ Batiment 440, 78352 Jouy en Josas Cedex, France

REFERENCE 2 (bases 1 to 54)
AUTHORS Cotinot, C., Kirszenbaum, M., Leonard, M., Gianguinto, L. and Vaiman, M.
TITLE Isolation of bovine Y-derived sequence: potential use in embryo sexing
JOURNAL Genomics 10 (3), 646-653 (1991)
MEDLINE 91365372
FEATURES
source 1. .54
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/chromosome="Y"
/map="short arm"
misc_feature <1..>54
BASE COUNT 14 a 18 c 11 g 11 t
ORIGIN

Query Match 43.5%; Score 17.4; DB 3; Length 54;
Best Local Similarity 77.8%; Pred. No. 3.8e+03;

Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 8 gacgtagatgagctacaagcgctgatt 34
||| ||||| ||| ||||| ||| |||
Db 42 GACCGAGATGTGCTCCACGAGAGTGT 16

RESULT 3
A62916/c 25 bp DNA PAT 12-MAR-1998
LOCUS A62916
DEFINITION Sequence 157 from Patent WO9719110.
ACCESSION A62916
VERSION A62916.1 GI:3716804
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 25)
AUTHORS Futreal, P.A., Wooster, R.F., Ashworth, A. and Stratton, M.R.
TITLE MATERIALS AND METHODS RELATING TO THE IDENTIFICATION AND SEQUENCING OF THE BRCA2 CANCER SUSCEPTIBILITY GENE AND USES THEREOF
JOURNAL Patent: WO 9719110-A 157 29-MAY-1997;
CANCER RES CAMPAIGN TECH (GB)
COMMENT Other publication AU 7635096 19970611
Other publication GB 2307477 19970528.
FEATURES
source 1. .25
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 7 a 8 c 2 g 8 t
ORIGIN

Query Match 41.5%; Score 16.6; DB 5; Length 25;
Best Local Similarity 82.6%; Pred. No. 7.8e+03;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 18 agctacaagcgctgattggat 40
|| | ||||| ||||| |||||
Db 24 AGTAGAAGGCGCTGATTGGAT 2

RESULT 4
I11638 60 bp DNA PAT 26-JUL-1995
LOCUS I11638
DEFINITION Sequence 12 from Patent US 5408041.
ACCESSION I11638
VERSION I11638.1 GI:909156
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 60)
AUTHORS Mundy, G.R., Gutierrez, G.E., Garrett, I.R., Sabatini, M., Izdicka, E., Burgess, W., Crumley, G.R., Morse, C.C. and Arnet, T.R.
TITLE Process of purifying antler-derived bone growth factors
JOURNAL Patent: US 5408041-A 12 18-APR-1995;
MEDLINE Location/Qualifiers
FEATURES
source 1. .60
/organism="unknown"
BASE COUNT 7 a 5 c 14 g 11 t 23 others
ORIGIN

Query Match 40.0%; Score 16; DB 5; Length 60;
Best Local Similarity 48.5%; Pred. No. 1.5e+04;
Matches 16; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY 8 gacgtagatgagctacaagcgctgattggat 40
||:|:| ||:|:| ||:|:| :|:|:|
Db 28 GAGTNGAIGCNCNTNCARTTYGINTGYGNGAY 60

LOCUS	Al0881	80 bp	DNA	PAT	29-SEP-1993
DEFINITION	Synthetic DNA (MCS) from patent EP0292763.				
ACCESSION	Al0881				
VERSION	Al0881.1	GI:490875			
KEYWORDS	synthetic construct.				
SOURCE	synthetic construct				
ORGANISM	artificial sequence.				
REFERENCE	1 (bases 1 to 80)				
AUTHORS	Crause,P.D., Hein,F., Jansen,H.W. and Uhlmann,E.				
TITLE	Gene-technological process for the preparation of angiogenines				
JOURNAL	Patent: EP 0292763-A 4 30-NOV-1988;				
FEATURES	HOECHST AKTIEGESELLSCHAFT				
SOURCE	location/Qualifiers				
	1..80				
	/organism="synthetic construct"				
	/db_xref="taxon:32630"				
BASE COUNT	17 a	27 c	20 g	16 t	
ORIGIN					
Query Match	40.0%;	Score 16;	DB 5;	Length 80;	
Best Local Similarity	62.5%;	Pred. No. 1.5e+04;			
Matches 25; Conservative	0;	Mismatches 15;	Indels	0;	Gaps 0;
1	gcacgtcgacgtagatgagctacaagcgctgattgggat	40			
79	gcattgcacgttgatgattgcattgcattgcacgaactgggctt	40			

[illegible]

ORGANISM	to mRNA, clone:YK19, YK45.			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;			
	Primates; Catarrhini; Hominoidea; Homo.			
REFERENCE	1 (bases 1 to 75)			
AUTHORS	Wang,L.			
JOURNAL	Direct Submission			
TITLE	Submitted (27-JUN-1995) to the DDBJ/EMBL/GenBank databases. L1			
REFERENCE	Wang, Japanese Red Cross Central Blood Center, Department of			
AUTHORS	Research; 4-1-31 Hiroo, Shibuya-Ku, Tokyo 150, Japan			
JOURNAL	(tel:03-5485-6009, Fax:03-3406-7892)			
TITLE	2 (bases 1 to 75)			
REFERENCE	Wang,L., Tadokoro,K., Tokunaga,K., Uchida,S., Mitsunaga,S.,			
AUTHORS	Bannai,M., Moriyama,S., Takai,K. and Juji,T.			
JOURNAL	Restricted usage of T-cell receptor Vb gene in post-transfusion			
TITLE	graft-versus-host disease			
REFERENCE	Unpublished (1995)			
AUTHORS	Location/Qualifiers			
JOURNAL	1. .75			
TITLE	/organism="Homo sapiens"			
REFERENCE	/isolate="YK"			
AUTHORS	/db_xref="taxon:9606"			
JOURNAL	/cell_type="T-lymphocyte"			
TITLE	/clone="YK19, YK45"			
REFERENCE	/sex="male"			
AUTHORS	/tissue_type="peripheral blood"			
JOURNAL	1. .75			
TITLE	/gene="TCRB"			
REFERENCE	<1. .9			
AUTHORS	/gene="TCRB"			
JOURNAL	/note="V14-segment"			
TITLE	<1. .>75			
REFERENCE	/gene="TCRB"			
AUTHORS	/codon_start=1			
JOURNAL	/product="T-cell receptor beta chain V14-D-J2.7-C2 region"			
TITLE	/protein_id="BA09405.1"			
REFERENCE	/db_xref="GI:1616882"			
AUTHORS	/translation="ASSSTGRTYEQYFGPGRILVTEDL"			
JOURNAL	10. .22			
TITLE	/gene="TCRB"			
REFERENCE	23. .66			
AUTHORS	/gene="TCRB"			
JOURNAL	/note="J2.7-segment"			
TITLE	67. .>75			
REFERENCE	/gene="TCRB"			
AUTHORS	/note="C2-segment"			
JOURNAL	BASE COUNT 18 a 23 c 25 g 9 t			
TITLE	ORIGIN			
REFERENCE	Query Match 39.0%; Score 15.6; DB 9; Length 75;			
AUTHORS	Best Local Similarity 63.2%; Pred. NO. 2.2e+04;			
JOURNAL	Matches 24; Conservative 0; Mismatches 14; Indels 0; Gaps 0;			
TITLE	Oy 1 gcaatgcagctagatgactacacagcgctgattgg 38			
REFERENCE	111 1111 11 11111 11 11 1111			
AUTHORS	Db 5 GCACGTCGACAGGAGAACCTACGACGACTCTCGCG 42			
JOURNAL	RESULT 8			
TITLE	128730/c 128730 35 bp DNA PAT 06-FEB-1997			
REFERENCE	LOCUS			
AUTHORS	DEFINITION Sequence 34 from patent US 5573925.			
JOURNAL	128730			
TITLE	ACCESSION			
REFERENCE	128730.1 GI:1819506			
AUTHORS	KEYWORDS			
JOURNAL	SOURCE			
TITLE	Unknown.			
REFERENCE	ORGANISM			
AUTHORS	Unclassified.			
JOURNAL	1 (bases 1 to 35)			
TITLE	Halazonetis,T.D.			
REFERENCE	p53 proteins with altered tetramerization domains			
AUTHORS	Patent: US 5573925-A 34 12-NOV-1996;			
JOURNAL				

Query Match 36.5%; Score 14.6; DB 5; Length 22;
Best Local Similarity 81.0%; Pred. No. 5.3e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 agctacaagcgctgattgg 38
|||||
Db 21 AGCTACAGGCACTCCTG 1

RESULT 13
LOCUS AR014522 33 bp DNA PAT 05-DEC-1998
DEFINITION Sequence 70 from patent US 5773582.
ACCESSION AR014522
VERSION AR014522.1 GI:3971976
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 33)
AUTHORS Shin,H., Shin,N., Lee,I. and Kang,S.
TITLE Tumor necrosis factor mutants
JOURNAL Patent: US 5773582-A 70 30-JUN-1998;
FEATURES Location/Qualifiers
source 1..33
/organism="unknown"

BASE COUNT 6 a 9 c 9 g 9 t
ORIGIN

Query Match 36.5%; Score 14.6; DB 5; Length 33;
Best Local Similarity 69.0%; Pred. No. 5.5e+04;
Matches 20; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 11 gtatgagctacaagcgctgattgg 39
|||||
Db 33 GTATGAGCAAGCGCCCTCTATTGCA 5

RESULT 14
LOCUS SYNANVAA 60 bp DNA SYN 27-APR-1993
DEFINITION Avian neovirus subgroup A DNA.
ACCESSION M60029
VERSION M60029.1 GI:207946
KEYWORDS
SOURCE Synthetic viral DNA.
ORGANISM Artificial construct.
REFERENCE 1 (bases 1 to 60)
AUTHORS Olsen,J.C., Bova-Hill,C., Grandgenett,D.P., Quinn,T.P.,
Manfredi,J.P. and Swanson,R.
TITLE Rearrangements in unintegrated retroviral DNA are complex and are
the result of multiple genetic determinants
JOURNAL J. Virol. 64, 5475-5484 (1990)
MEDLINE 91012792
FEATURES Location/Qualifiers
source 1..60
/organism="synthetic construct"
/db_xref="taxon:32530"

BASE COUNT 16 a 15 c 17 g 12 t
LTR
ORIGIN

Query Match 36.5%; Score 14.6; DB 14; Length 60;
Best Local Similarity 69.0%; Pred. No. 5.8e+04;
Matches 20; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9 acgtatgagctacaagcgctgattgg 37
|||||
Db 15 ACCTGCATGAAGCAAGGCTTCATTGG 43

RESULT 15
LOCUS JS3PRSEC 62 bp DNA BCT 26-APR-1993
DEFINITION plasmid pJS37 promoter region DNA.
ACCESSION M34566
VERSION M34566.1 GI:149149
KEYWORDS
SOURCE Plasmid pJS37 DNA, from E.coli (strain C600).
ORGANISM plasmid pJS37

REFERENCE 1 (bases 1 to 62)
AUTHORS Ballester,S., Alonso,J.C., Lopez,P. and Espinosa,M.
TITLE Comparative expression of the pC194 cat gene in Streptococcus
pneumoniae, Bacillus subtilis and Escherichia coli
JOURNAL Gene 86, 71-79 (1990)
MEDLINE 90185250
FEATURES Location/Qualifiers
source 1..62
/organism="plasmid pJS37"
/plasmid="plasmid pJS37"
/db_xref="taxon:2582"

misc_signal 21..43
-35_signal 21..26
-10_signal 43..48
/note="PT -35 region"
/note="PT -10 region"
mRNA 52..>62
/note="cat mRNA"

BASE COUNT 23 a 3 c 14 g 22 t
ORIGIN

Query Match 36.5%; Score 14.6; DB 1; Length 62;
Best Local Similarity 69.0%; Pred. No. 5.8e+04;
Matches 20; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 12 taatgagctacaagcgctgattgg 40
|||||
Db 34 TATATGCTGAATAATGTTTAATTGGAT 62

Search completed: June 4, 2000, 16:05:46
Job time: 27898 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:24:21 ; Search time 322.35 Seconds
(without alignments)
31.046 Million cell updates/sec

Title: US-09-164-714-20

Perfect score: 40
Sequence: 1 gcatgtcgacgtagatgactcaagaagcgtattggat 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 431286

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.6	46.5	88	1	V39983 Adeno-associated v
2	17.4	43.5	49	1	N70433 Probe b.c.1.2. New
3	17	42.5	40	1	V47799 Maize polypomorphic
4	17	42.5	53	1	V44499 Biotinylated seque
5	16.8	42.0	98	1	V07390 Synthetic Plasmid
6	16.6	41.5	25	1	T92588 BRCA2 cancer susce
7	16.6	41.5	50	1	X18007 primer HSA2 for ge
8	16.6	41.5	98	1	V07389 Synthetic Plasmid
9	16.4	41.0	52	1	T73522 primer RH104 for m
10	16.4	41.0	52	1	T73541 primer RH104 for m
11	16	40.0	60	1	Q49524 Antlerin-2 (OT-2)
12	15.8	39.5	96	1	T49274 HVRI region of E2
13	15.6	39.0	39	1	X05915 HSV VP16 minimal a
14	15.6	39.0	79	1	X05917 HSV VP16 minimal a
15	15.4	38.5	30	1	V12825 primer for yeast 5
16	15.2	38.0	35	1	T32845 Tumour suppressor
17	15	37.5	96	1	T49275 HVRI region of E2
18	14.8	37.0	38	1	N93056 Oligonucleotide fo
19	14.8	37.0	44	1	V57479 Human apoptosis in
20	14.8	37.0	71	1	T89124 Probe #4 amplified
21	14.8	37.0	88	1	O44252 Probe #4 amplified
22	14.8	37.0	88	1	O44252 VHL disease gene p
23	14.6	36.5	22	1	Q78903 primer MU-1 for vo
24	14.6	36.5	22	1	T90934 Oligonucleotide sr
25	14.6	36.5	35	1	X15164 PCR primer #1. Scr
26	14.6	36.5	50	1	X02203 staphylococcus aur
27	14.6	36.5	75	1	V76210 C-C chemokine rece
28	14.4	36.0	25	1	V07412 Scal methylase gen
29	14.4	36.0	45	1	T70165 Bacillus thuringie
30	14.4	36.0	46	1	Q68522 ctyic gene primer
31	14.4	36.0	48	1	T09117 PCMV-Fab105-Protam
32	14.4	36.0	48	1	O99858 53 bp intronic seq
33	14.4	36.0	53	1	V34508 Sequence of human
34	14.4	36.0	80	1	O20226

35	14.4	36.0	80	1	O08622 HPV-16 capture o11
36	14.4	36.0	84	1	X32350 Artificial NS4 mos
37	14.2	35.5	24	1	V27907 Rat GALR3 third tr
38	14.2	35.5	65	1	T39537 Nucleotide used in
39	14.2	35.5	68	1	V52996 Oligonucleotide us
40	14.2	35.5	68	1	V52997 Oligonucleotide us
41	14.2	35.5	99	1	V42880 RNA aptamer bindin
42	14.2	35.5	100	1	T30900 primer 30 for 95 k
43	14	35.0	33	1	Q84735 Primer to clone th
44	14	35.0	33	1	T03317 Murine IgG heavy c
45	14	35.0	35	1	Q36177 Mutagenic primer C

ALIGNMENTS

RESULT 1	
V39983/C	
ID V39983 standard; DNA; 88 BP.	
AC V39983;	
DT 05-OCT-1998 (first entry)	
DE Adeno-associated virus cap gene pSub201 primer binding region.	
KM Adeno-associated virus; AAV; adenovirus; vector-mediated delivery;	
KW liver; viral particle; tumour; expression; blood disease; haemophilia;	
KM metabolic disease; familial hypercholesterolaemia; primer; ss.	
OS Synthetic.	
PN W09824479-A1.	
PD 11-JUN-1998.	
PF 02-DEC-1997; U21398.	
PE 25-JUN-1997; US-882044.	
PR 02-DEC-1996; US-032506.	
PA (SOMA-) SOMATIX THERAPY CORP.	
PA (UNIW) UNIV WASHINGTON.	
PI Cohen L, Danos O, Kay M, Snyder R, Thompson AR;	
DR WPI; 98-333055/29.	
PT Expression of polynucleotide(s) in mammals - by administering viral	
PT particles comprising recombinant adeno-associated virus to liver	
PT cells; used for; e.g. treating liver tumours	
PS Example 7; Fig 5; 63pp; English.	
CC Methods have been developed for: (1) expressing a polynucleotide (PN) in	
CC a mammal comprising administering viral particles comprising a	
CC recombinant adeno-associated virus (RAV) vector to liver cells of the	
CC mammal, where the RAV vector comprises the PN, such that the PN, or a	
CC portion, is expressed in the mammal; (2) treating a liver disease or	
CC disorder in a mammal, comprising administering a RAV vector to the	
CC liver cells of the mammal; (3) determining the presence of wild-type AAV	
CC and infectious AAV generated by recombination of a helper AAV and a	
CC vector AAV containing a transgene in a sample of RAVV, where the vector	
CC AAV contains nucleotide sequences or has an order of nt sequences	
CC different from that of wild-type AAV. The RAVV vectors may be used in the	
CC preparation of therapeutics useful for treating a disease or disorder in	
CC a mammal or for gene therapy. The methods permit expression of diffusible	
CC polypeptides in the liver which provides access to the circulation and	
CC permits systemic delivery of therapeutic proteins and polypeptides. They	
CC can be used for the treatment of blood diseases, e.g. haemophilia,	
CC metabolic diseases, e.g. familial hypercholesterolaemia, liver specific	
CC diseases, e.g. glycogen storage diseases or liver tumours. The present	
CC sequence represents an oligonucleotide from the present invention.	
SQ Sequence 88 BP; 22 A; 22 C; 20 G; 24 T;	
Query Match 46.5%; Score 18.6; DB 1; Length 88;	
Best local similarity 72.7%; Pred. No. 36;	
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;	
OY 1 gcatgtcgacgtagatgactcaagaagcgtat 33	
DB 82 GCATGCTACGTAGATGACTCAAGCGGTT 50	
RESULT 2	
N70433	
ID N70433 standard; DNA; 49 BP.	

AC N70433:
16-FEB-1991 (first entry)
DE Probe b.c.1.2.
KM Probe b.c.1.2.; male-specific probe; ruminant; spermatazoa;
OS Bos taurus.
PN EP-235046-A.
PD 02-SEP-1987.
PF 27-FEB-1987; 400434.
PR 28-FEB-1986; FR-002811.
PR 09-SEP-1986; FR-012616.
PA (INRG) INST NAT RECH AGRON.
PA (INSP) INST PASTEUR.
PA (COMS) COMMISSARIAT A L'ENERGIE ATOMIQUE.
PI Bishop C, Colinet C, Fellous M, Kirszenbau M, Vainan M;
DR WPI: 87-244362/35.
PT New DNA hybridisation probes - specific to ruminant male genome
CC Disclosure: page 5; 20pp: french.
CC The probe is specific for the male genome of ruminants (esp.
CC cattle). It has a hybridisation profile, as determined by
CC hybridisation with EcoRI-digested male genomic DNA, with at least
CC one band specific to the male genome of the ruminant. The probe is
CC repeated at least 2000 times in the bovine male genome. The probe
CC is useful for sexing ruminant embryos or foetuses, or for
CC monitoring the proportion of spermatozoa carrying the Y chromosome,
CC esp. for separating X- or Y-bearing spermatozoa.
SQ Sequence 49 BP; 10 A; 10 C; 16 G; 13 T;

Query Match 43.5%; Score 17.4; DB 1; Length 49;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 8 gacgtagatgagctacaaggcgtgatt 34
||| ||||| ||| ||||| ||| ||
DB 12 GACCGAGATGAGTCTCCACAGAGTGT 38

RESULT 3
V47799/c
ID V47799 standard; DNA: 40 BP.
AC V47799:
DT 14-OCT-1998 (first entry)
DE Maize polymorphic site oligonucleotide marker Wx1-G2/G6-5.
KM Maize; marker: probe; PCR primer; polymorphism; vegetal sequence;
KW polymorphic site; corn; graminæ species; ss.
OS Synthetic.
OS zea sp.
PN WO9830717-A2.
PD 16-JUL-1998.
PF 02-DEC-1997; E07134.
PR 02-DEC-1996; US-032069.
PA (BIOC-) BIOCEM SA.
PI Murigneux A;
DR WPI: 98-399160/34.
PT Vegetal sequences including single nucleotide polymorphism - useful,
PT e.g. to determine polymorphisms in plants, determine strain in plant
PT breeding and to correlate polymorphisms with phenotypic traits
PT Claim 2; Page 13; 32pp: English.
CC The present invention describes a nucleic acid segment comprising at
CC least 10 contiguous nucleotides from a vegetal acid segment including a
CC polymorphic site which is a single nucleotide polymorphism (SNP), or the
CC complement of the segment. Also described are: (1) an allele-specific
CC oligonucleotides hybridising to segment, or their complements, and (2) a
CC method of analysing nucleic acids from a subject, by determining if a
CC base is occupying any one (or a set) of polymorphic sites in 261
CC sequences derived from six maize lines (see V47701 to V47961). The
CC segments are useful in fingerprint analysis in plants to determine which
CC polymorphisms are present, which strain a plant belongs to and to
CC distinguish between strains. The polymorphisms may correlate with
CC phenotypic traits (e.g. plant growth rate or crop yield), and the
CC segments are useful to determine the presence/absence of specific
CC polymorphisms correlating with the existence/absence of particular
CC traits. The segments are also useful in marker assisted back-cross

CC techniques to select plants with a higher percentage of recurrent parent
CC in a back-cross population. Segments incorporate SNPs which occur more
CC frequently than other polymorphism types and are therefore more likely to
CC be located close to genetic loci of interest; different forms of
CC characterised SNPs are also often easier to detect than other
CC polymorphism types.
SQ Sequence 40 BP; 7 A; 17 C; 9 G; 7 T;

Query Match 42.5%; Score 17; DB 1; Length 40;
Best Local Similarity 69.7%; Pred. No. 1.4e+02;
Matches 23; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 8 gacgtagatgagctacaaggcgtgattgatt 40
||| ||||| ||||| ||||| |||||
DB 35 GCCCGAGATGAGTCTCTCGCGCTAGTAGGGCCT 3

RESULT 4
V44499/c
ID V44499 standard; DNA: 53 BP.
AC V44499:
DT 20-OCT-1998 (first entry)
DE Biotinylated sequencing template.
KM DNA sequencing; heterogeneous population; primer; ss.
OS Synthetic.
OS WO9831831-A1.
PN 23-JUL-1998.
PD 15-JAN-1998; G00130.
PR 15-JAN-1997; GB-000760.
PA (BRAX-) BRAX GENOMICS LTD.
PI Schmidt G, Thompson AH;
DR WPI: 98-427563/36.
PT Sequencing of DNA - using an array of hybridisable probes comprising
PT a label cleavably attached to a known sequence of predetermined
PT length
PT Disclosure: Page 25; 51pp: English.
PS This sequence is used in a method for sequencing DNA. The method allows
CC a heterogeneous population of nucleic acid fragments to be sequenced
CC simultaneously. The method does not require traditional gel methods to
CC acquire sequence information. Since the entire process takes place in
CC solution and is an iterative process, the steps involved could be
CC performed by a liquid-handling robot.
SQ Sequence 53 BP; 17 A; 8 C; 12 G; 16 T;

Query Match 42.5%; Score 17; DB 1; Length 53;
Best Local Similarity 69.7%; Pred. No. 1.5e+02;
Matches 23; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 3 atgtcagctagatgagctacaaggcgtgatt 35
||||| ||| ||| ||| ||| |||||
DB 34 ATGTCTACTTACTTAACATCATCAGATCGATT 2

RESULT 5
V07390
ID V07390 standard; DNA: 98 BP.
AC V07390:
DT 13-OCT-1998 (first entry)
DE Synthetic plasmid construction oligonucleotide.
KM Mutant; glucocorticoid receptor; plasmid; constitutively active;
KW GR; gene expression; regulation; gene therapy; steroid; treatment;
OS Ashtma; arthritis; oligonucleotide; ss.
OS Synthetic.
PN WO9818925-A2.
PD 07-MAY-1998.
PF 28-OCT-1997; U19607.
PR 29-OCT-1996; US-029964.
PA (BAYU) BAYIOR COLLEGE MEDICINE.
PA (GENE-) GENEMEDICINE INC.
PI Kittle JD, Ledebur HC, Omalley B, Tsai M, Tsai SY,
Wang Y;

DR WPI: 98-272220/24.
PT Modified glucocorticoid receptors able to bind non-natural ligands -
PT for regulating gene expression in gene therapy without requiring
PT exogenous steroid(s), particularly for treatment of arthritis and
PT asthma
PS Example 17: Page 70; 124pp; English.
CC The sequence is that of an oligonucleotide which was used in the
CC construction of a plasmid expressing a constitutively active mutant
CC glucocorticoid receptor (GR) protein.
SQ Sequence 98 BP; 23 A; 26 C; 26 G; 23 T;

Query Match 42.0%; Score 16.8; DB 1; Length 98;
Best Local Similarity 66.7%; Pred. No. 2e+02; Mismatches 12; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 5 gtcacgtagatgagctacaaagcgtagttggat 40
| | | | | | | | | | | | | | | | | |
Db 1 GGCCGGTTATATCTAGAACGGCTGATTCGGAT 36

RESULT 6
T92588/c
ID T92588 standard; DNA; 25 BP.
AC T92588:
DT 04-FEB-1998 (first entry)
DE BRCA2 cancer susceptibility gene exon 25B PCR primer F for SSCP.
KW BRCA2 cancer susceptibility gene; breast cancer; ovarian cancer;
KW gene therapy; prostate cancer; colorectal cancer; ocular melanoma;
KW leukemia; human; single stranded conformation polymorphism test;
KW SSCP; PCR primer; ss.
OS Synthetic.
OS Homo sapiens.
PN GB2307477-A.
PD 28-MAY-1997.
PF 25-NOV-1996; 024453.
PR 28-AUG-1996; GB-017961.
PR 23-NOV-1995; GB-023959.
PR 14-DEC-1995; GB-023535.
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (UYDU-) UNIV DUKE.
PI Ashworth A, Futreal PA, Stratton MR, Wooster RF;
DR WPI: 97-261854/24.
PT Nucleic acid molecules comprising part or all of the BRCA2 cancer
PT susceptibility gene - useful for diagnosis, prognosis or therapeutic
PT treatment of cancer.
PS Example 1; Fig 8; 124pp; English.
CC The present sequence represents a PCR primer for single stranded
CC conformation polymorphism testing of the BRCA2 cancer susceptibility
CC gene. The nucleic acid molecule can be used to construct probes for
CC screening cDNA or genomic libraries, sequencing positive clones
CC obtained, and assembling the full length BRCA2 sequence. The BRCA2
CC nucleic acid molecules and proteins are useful in a method of medical
CC treatment, preferably gene therapy, especially for treating cancer,
CC where the cancer is female or male breast cancer, ovarian, prostate or
CC colorectal cancer, ocular melanoma or leukaemia. In particular
CC antisense oligonucleotides capable of hybridising to the BRCA2 nucleic
CC acid, pre-mRNA or mature mRNA are used so that the expression of the
CC BRCA2 nucleic acid is reduced or prevented. The nucleic acid molecules
CC are also useful in a method for diagnosing susceptibility or
CC predisposition to cancer in a patient. The nucleic acid molecules are
CC used to design probes or primers for PCR to determine or detect the
CC presence of mutations in a sample of nucleic acid from a patient. The
CC BRCA2 promoter region is useful for screening for substances which
CC modulate the expression of nucleic acid under control of the promoter.
CC Antibodies are used to determine the presence, amount or location in a
CC cell of a BRCA2 polypeptide or its mutant forms. The polypeptides are
CC used to screen for binding partners, these are useful to screen for
CC substances which mimic the activity of BRCA2 polypeptide, which can be
CC used as cancer therapeutics.
SQ Sequence 25 BP; 7 A; 8 C; 2 G; 8 T;

Query Match 41.5%; Score 16.6; DB 1; Length 25;
Best Local Similarity 82.6%; Pred. No. 1.9e+02; Mismatches 4; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 18 agctacaaagcgtagttggat 40
| | | | | | | | | | | | | | | | | |
Db 24 AGTAAGAAGCGCTGATTTGGCAT 2

RESULT 7
X18007
ID X18007 standard; DNA; 50 BP.
AC X18007:
DT 21-MAY-1999 (first entry)
DE Primer HSA2 for generating plasmid pAD/L420-HSA.
KW Recombinant; adenovirus; vector; interleukin; primer; PCR; amplification;
KW cancer; cytostatic activity; tumour; ss.
OS Synthetic.
OS Mus sp.
PN W09902685-A1.
PD 21-JAN-1999.
PF 13-JUL-1998; NL0406.
PR 11-JUL-1997; EP-202167.
PA (INTR-) INTROGENE BV.
PI Bout A, Draijer-Van Der Kaaden ME;
DR WPI: 99-120886/10.
PT Treating tumours with a composition having interleukin-3 activity -
PT particularly using a recombinant adenoviral vector encoding that
PT activity
PS Example 2; Page 26; 79pp; English.
CC The invention relates to a pharmaceutical composition, comprising a
CC recombinant adenoviral vector encoding interleukin-3 (IL-3) activity.
CC The primers X21500 and X18007 were used to construct the recombinant
CC plasmid pAD/L420-HSA by PCR amplifying the murine HSA gene.
CC The recombinant adenoviral vector encoding IL-3 activity can be used
CC in the manufacture of a pharmaceutical composition for the systemic
CC treatment of cancer. A composition comprising IL-3 activity is used
CC with a composition comprising cytostatic activity, to treat tumours.
SQ Sequence 50 BP; 17 A; 8 C; 11 G; 14 T;

Query Match 41.5%; Score 16.6; DB 1; Length 50;
Best Local Similarity 64.1%; Pred. No. 2.1e+02; Mismatches 14; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 gcatgcacgtagatgagctacaaagcgtagttggga 39
| | | | | | | | | | | | | | | | | |
Db 12 GCTGTGCACATCGATCTACTACAGTAGAGATGTAGAA 50

RESULT 8
V07389/c
ID V07389 standard; DNA; 98 BP.
AC V07389:
DT 13-OCT-1998 (first entry)
DE Synthetic plasmid construction oligonucleotide.
KW Mutant; glucocorticoid receptor; plasmid; constitutively active;
KW GR; gene expression; regulation; gene therapy; steroid; treatment;
KW asthma; arthritis; oligonucleotide; ss.
OS Synthetic.
OS W09818925-A2.
PN W09818925-A2.
PD 07-MAY-1998.
PF 28-OCT-1997; U19607.
PR 29-OCT-1996; US-029964.
PA (BAYU) BAYLOR COLLEGE MEDICINE.
PA (GENE-) GENEMEDICINE INC.
PI Kittle JD, Ledebur HC, Omalley B, Tsai M, Tsai SY,
PI Wang Y;
DR WPI: 98-272220/24.
PT Modified glucocorticoid receptors able to bind non-natural ligands -
PT for regulating gene expression in gene therapy without requiring
PT exogenous steroid(s), particularly for treatment of arthritis and
PT asthma

PI Izlicka E, Morse CC, Mundy GR, Sabatini M;
DR WPI: 93-320678/40.
DR P-PSDB; R41735.
PT Antler-derived bone growth factors - useful to promote growth and
PT differentiation of osteoblastic cells
PS Example 14; Page 34; 98pp; English.
CC The purified "antlerin" OR-2 was isolated from deer antler
CC deer antler tissue by sequential use of EDTA extraction, affinity
CC chromatography on heparin sepharose resin, gel permeation
CC Chromatography on Sephacryl S-100 resin and HPLC chromatography on
CC C18, C4 and C8 resins. OR-2 stimulates growth of human cells with
CC the osteoblast phenotype. Q49524 is a degenerate coding sequence
CC for OR-2.
SQ Sequence 60 BP; 7 A; 5 C; 14 G; 11 T;

Query Match 40.0%; Score 16; DB 1; Length 60;
Best Local Similarity 48.5%; Pred. No. 3.9e+02;
Matches 16; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 8 gaagtagatgagctacaagcgtagattggat 40
||:||||:||||:||||:||||:
DB 28 GAYGTNGAYGNCNTCARTYGTNTGYGNAY 60

RESULT 12
T49274/C
ID T49274 standard; DNA; 96 BP.
AC T49274;
DT 08-SEP-1997 (first entry)
DE HVRI region of E2 gene of HCV isolate SA6, subtype 5a, genotype 5.
KW Hypervariable region 1; HVRI, E2 gene; hepatitis C virus; HCV;
KW Polypeptide; hydrophilic domain; antigenic determinant; epitope;
KW viral life cycle; diagnosis; prevention; treatment; HCV infection;
KW vaccine; mammal; ss.
OS Hepatitis C virus.
PN W09640764-A2.
PD 19-DEC-1996.
PE 05-JUN-1996; U09340.
PR 07-JUN-1995; US-484322.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Bukh J, Miller RH, Purcell RH;
DR WPI: 97-052230/05.
DR P-PSDB; W08791.
PT New isolated hepatitis C virus sequences - correspond to
PT hyper-variable region 1 of the envelope 2 gene of different HCV
PT isolates
PS Claim 1; Page 52; 101pp; English.
CC The sequences given in T49231-79 represent nucleic acids derived from
CC hyper-variable region 1 (HVRI) of the E2 gene of 49 hepatitis C virus
CC (HCV) isolates. HVRI is defined as starting at amino acid 384 of the
CC HCV polypeptide. The HVRI amino acid sequences typically contain 3
CC hydrophilic domains which are located at the N- and C-termini and
CC roughly in the middle of HVRI. These hydrophilic domains may represent
CC important antigenic determinants and the C-terminal hydrophilic domain
CC is universally conserved. This conservation may indicate that this
CC region is not only important as an immunodominant epitope for HCV but
CC can also play an important part in the viral life cycle. These peptides
CC can be used to develop products for the diagnosis, prevention and
CC treatment of HCV infection. The peptide sequences and vectors containing
CC the DNAs can be used in the preparation of a vaccine to immunise a
CC mammal against HCV infection.
SQ Sequence 96 BP; 22 A; 31 C; 23 G; 20 T;

Query Match 39.5%; Score 15.8; DB 1; Length 96;
Best Local Similarity 74.1%; Pred. No. 5.2e+02;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 11 gtatgatgactacaagcgtagattgg 37
||||| ||||| ||||| ||||| |||||
DB 96 GTATGTAGCTGCAAGTTCCTTCGG 70

RESULT 13
X05915/C
ID X05915 standard; DNA; 39 BP.
AC X05915;
DT 07-MAY-1999 (first entry)
DE HSV VP16 minimal activation domain encoding synthetic oligo [F].
KW Transcription activator; fusion protein; transcription domain; HSV VP16;
KW mutated; acidic region; herpes simplex virus; vitron protein 16; cancer;
KW gene therapy; genetic disease; rheumatoid arthritis; hypopituitarism;
KW benign prostatic hypertrophy; viral disease; transactivator; ss.
OS Synthetic.
PN W09901549-A1.
PD 14-JAN-1999.
PE 01-JUL-1998; U13993.
PR 03-JUL-1997; US-888080.
PA (BARO/) BARON U.
PA (BADI) BASF AG.
PI Baron U, Bujard H, Gossen M;
DR WPI: 99-106051/09.
PT New transcriptional activator (transactivators) fusion proteins -
PT contain a mutated acidic region of herpes simplex virus vitron
PT protein 16, useful for, e.g. regulating particular gene expression
PS Examples; Page 30; 56pp; English.
CC The invention relates to a fusion protein that activates transcription.
CC The protein comprises a DNA binding domain operably linked to a
CC transcriptional activation domain, comprising a mutated acidic region of
CC herpes simplex virus vitron protein 16 (HSV VP16), residues 436-447 of
CC HSV VP16 (the mutation is a substitution at position 442 as compared to
CC wild type HSV VP16). Nucleic acid molecules encoding the fusion proteins,
CC the polypeptides and host cells containing a recombinant vector
CC comprising the nucleic acid form kits that are useful for regulating the
CC expression of desired genes in a fast, efficient and controlled manner.
CC The kits can be applied to the study of cellular development and
CC differentiation in eukaryotic cells, plants and animals. In addition to
CC regulating the expression of site-specific recombinases to allow for
CC irreversible modification of the genotype of a transgenic organism at a
CC certain stage of development. The kits are also useful for gene therapy
CC for treating genetic or acquired diseases, especially cancer, rheumatoid
CC arthritis, hypopituitarism, benign prostatic hypertrophy and viral
CC diseases. The kits are useful for in vitro and in vivo production of
CC proteins. The polypeptides are useful for stimulating expression of
CC specific genes in animals to mimic the pathophysiology of human disease
CC to create animal models of human disease. The transactivator polypeptides
CC are tolerated at high concentrations in cells (squenching is avoided).
CC which enables HSV VP16 to regulate levels of gene expression in a very
CC precise manner. Sequences X05915-919 represent synthetic oligos encoding
CC the minimal activation domains of HSV VP16.
SQ Sequence 39 BP; 6 A; 15 C; 12 G; 6 T;

Query Match 39.0%; Score 15.6; DB 1; Length 39;
Best Local Similarity 70.0%; Pred. No. 5.3e+02;
Matches 21; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 gcatgcagctagatgagctacaagcgt 30
||||| ||||| ||||| ||||| |||||
DB 37 GCATGTCAGGTCGACAGTCTCAGGCGCT 8

RESULT 14
X05917/C
ID X05917 standard; DNA; 79 BP.
AC X05917;
DT 07-MAY-1999 (first entry)
DE HSV VP16 minimal activation domain encoding synthetic oligo [FG].
KW Transcription activator; fusion protein; transcription domain; HSV VP16;
KW mutated; acidic region; herpes simplex virus; vitron protein 16; cancer;
KW gene therapy; genetic disease; rheumatoid arthritis; hypopituitarism;
KW benign prostatic hypertrophy; viral disease; transactivator; ss.
OS Synthetic.
PN W09901549-A1.
PD 14-JAN-1999.
PE 01-JUL-1998; U13993.
PR 03-JUL-1997; US-888080.
PA (BARO/) BARON U.
PA (BADI) BASF AG.
PI Baron U, Bujard H, Gossen M;
DR WPI: 99-106051/09.
PT New transcriptional activator (transactivators) fusion proteins -
PT contain a mutated acidic region of herpes simplex virus vitron
PT protein 16, useful for, e.g. regulating particular gene expression
PS Examples; Page 30; 56pp; English.
CC The invention relates to a fusion protein that activates transcription.
CC The protein comprises a DNA binding domain operably linked to a
CC transcriptional activation domain, comprising a mutated acidic region of
CC herpes simplex virus vitron protein 16 (HSV VP16), residues 436-447 of
CC HSV VP16 (the mutation is a substitution at position 442 as compared to
CC wild type HSV VP16). Nucleic acid molecules encoding the fusion proteins,
CC the polypeptides and host cells containing a recombinant vector
CC comprising the nucleic acid form kits that are useful for regulating the
CC expression of desired genes in a fast, efficient and controlled manner.
CC The kits can be applied to the study of cellular development and
CC differentiation in eukaryotic cells, plants and animals. In addition to
CC regulating the expression of site-specific recombinases to allow for
CC irreversible modification of the genotype of a transgenic organism at a
CC certain stage of development. The kits are also useful for gene therapy
CC for treating genetic or acquired diseases, especially cancer, rheumatoid
CC arthritis, hypopituitarism, benign prostatic hypertrophy and viral
CC diseases. The kits are useful for in vitro and in vivo production of
CC proteins. The polypeptides are useful for stimulating expression of
CC specific genes in animals to mimic the pathophysiology of human disease
CC to create animal models of human disease. The transactivator polypeptides
CC are tolerated at high concentrations in cells (squenching is avoided).
CC which enables HSV VP16 to regulate levels of gene expression in a very
CC precise manner. Sequences X05915-919 represent synthetic oligos encoding
CC the minimal activation domains of HSV VP16.
SQ Sequence 39 BP; 6 A; 15 C; 12 G; 6 T;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 13:53:37 ; Search time 4521.53 Seconds
(without alignments)
35.857 Million cell updates/sec

Title: US-09-164-714-20

Perfect score: 40
Sequence: 1 gcatgtcgaactgagatgagctacaagcgatgattggat 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 119968

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_est48:*
83: gb_est49:*
84: gb_est50:*
85: gb_est51:*
86: em_est35:*
87: em_est36:*
88: em_est37:*
89: em_est38:*
90: gb_est52:*
91: gb_est53:*
92: gb_est54:*
93: gb_est55:*
94: gb_est56:*
95: em_est39:*
96: em_est40:*
97: em_est41:*
98: em_est42:*
99: em_est43:*
100: em_est44:*
101: em_est45:*
102: gb_est57:*
103: gb_est58:*
104: em_est46:*
105: gb_est59:*
106: gb_est60:*
107: gb_est61:*
108: gb_est62:*
109: gb_est63:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	17.2	43.0	68	50	F26502	F26502 HSPD14000 H
2	17	42.5	94	40	AA967015	AA967015 ua3b11.r
3	16.4	41.0	77	37	AA730273	AA730273 n294b05.s
4	16.2	40.5	86	33	AA431534	AA431534 zw7ch06.r
5	16.2	40.5	94	36	AA641353	AA641353 nr91e01.s
6	16	40.0	66	40	AA912094	AA912094 O157g09.s
7	16	40.0	82	24	H66370	H66370 yu14b03.r1
8	16	40.0	82	30	AA212741	AA212741 mw83c10.r
9	15.8	39.5	58	23	H49949	H49949 y025c03.r1
10	15.8	39.5	85	59	AT767746	AT767746 wh3c07.x
11	15.8	39.5	96	23	R84750	R84750 yq28a09.r1
12	15.8	39.5	100	26	w77250	w77250 me65e02.r1
13	15.6	39.0	71	45	AT360727	AT360727 qx98d03.x
14	15.4	38.5	67	42	AT159257	AT159257 vz88b08.r
15	15.4	38.5	70	74	AM215038	AM215038 up03h04.y
16	15.4	38.5	98	50	AT689524	AT689524 tx13d01.x
17	15.2	38.0	78	44	AT267735	AT267735 ap64a02.x
18	15.2	38.0	94	27	AA000065	AA000065 mg27g09.r
19	15.2	38.0	95	49	AT664297	AT664297 ue84a12.r
20	15.2	38.0	99	62	AT903060	AT903060 OV-BT022-774264.yc56g05.s1
21	15	37.5	52	21	T74264	T74264 y056g05.s1
22	15	37.5	52	37	AA723687	AA723687 ah85c05.s
23	15	37.5	82	44	AT324951	AT324951 ml174h09.x
24	15	37.5	84	42	AF027883	AF027883 AF027883
25	15	37.5	88	21	T99680	T99680 ye67d08.r1
26	15	37.5	94	38	AA765118	AA765118 nz61c10.s
27	15	37.5	97	48	AT656392	AT656392 tt50e11.x
28	15	37.5	97	62	AT928344	AT928344 wo44g01.x
29	15	37.5	100	81	AAW53970	AAW53970 zeh10745
30	14.6	36.5	55	30	AA214324	AA214324 zq91e11.s
31	14.6	36.5	73	40	AA936694	AA936694 o179d04.s
32	14.6	36.5	73	24	D18181	D18181 MUGS00446
33	14.6	36.5	77	45	AT393278	AT393278 tg14e07.x
34	14.6	36.5	80	20	T11114	T11114 hbc700.Huma
35	14.6	36.5	89	30	AA266257	AA266257 mz67a02.r
36	14.6	36.5	91	27	AA039440	AA039440 zk39b09.r
37	14.6	36.5	92	42	AT121560	AT121560 uc43c09.r
38	14.6	36.5	95	25	U44258	U44258 ENM44258.As
39	14.6	36.5	100	64	AA063099	AA063099 SMAMC443
40	14.4	36.0	55	39	AA862559	AA862559 oh44c03.s
41	14.4	36.0	58	63	AT940991	AT940991 sb82h08.y
42	14.4	36.0	70	26	W47924	W47924 mc84e09.r1
43	14.4	36.0	82	39	AA845785	AA845785 ak77h10.s
44	14.4	36.0	87	73	AM168701	AM168701 x194e06.x
45	14.4	36.0	88	37	AA699901	AA699901 zi61e02.s

ALIGNMENTS

RESULT 1
LOCUS F26502 68 bp mRNA EST 13-MAY-1999
DEFINITION HSPD14000 HM3 Homo sapiens cDNA clone s4000056B08, mRNA sequence.
ACCESSION F26502
VERSION F26502.1 GI:4812128
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 68)
Pandofo,D., Topo,S., Trevisan,S., Scarso,S. and Valle,G.
Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization
Genome Res. 6 (1), 35-42 (1996)

AUTHORS

TITLE

JOURNAL MEDLINE
96276048
On Mar 10, 1998 this sequence version replaced gi:2948143.
Contact: Valle G.

CRIBI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
http://grp.bio.unipd.it.
Location/Qualifiers

FEATURES
source

1. 68
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s4000056B08"
/clone_lib="HM3"
/sex="female"

/tissue_type="pectoral muscle (after mastectomy)"
/note="Vector: pCDNA1 (Invitrogen); Site_1: BstXI;
Site_2: NotI; The library was constructed by G.
Lanfranchi. This library is not subtracted nor normalized.

The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-Diotin-AACCGGCTCGAGCGCCCTTTTCTTTTCTTTTCTTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pCDNA1 vector."

BASE COUNT 22 a 19 c 18 g 9 t
ORIGIN

Query Match 43.0%; Score 17.2; DB 50; Length 68;
Best Local Similarity 73.3%; Pred. No. Jc+03; 8; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 2 catgtcagctagatgacacagcgctg 31
Db 2 CAAGGCCATGACAGATGTACAGCGCTG 31

RESULT 2
LOCUS AA967015/c 94 bp mRNA EST 19-MAY-1998
DEFINITION ua3b11.r1 Soares_mammary_gland_NBMG Mus musculus cDNA clone
IMAGE:1349085 5' similar to SW:THIO.CYAC P37395 THIOREDOXIN.;
MUS sequence.
ACCESSION AA967015
VERSION AA967015
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 94)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HM1 Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Jan 14, 1998 this sequence version replaced gi:1796941.
Contact: Marra M/Mouse EST Project
WashU-HM1 Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LUNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:697877

Trace considered overall poor quality

Possible reversed clone: Similarity on wrong strand

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 1.

[illegible]

```

/clone="IMAGE:1303089"
/clone_lib="NCI_CGAP_GCB1"
/issue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-GTGACCAATCGAAGGAGGAGGCGCCGCTCAATTTTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pRT3D vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT      25 a      6 c      17 g      29 t

ORIGIN

Query Match
Best Local Similarity 76.9%; Pred. No. 6.4e+03;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

0Y 15 atgagctcaagcgctgattggat 40
      |||| | |||| ||||| ||
Db 30 ATGACTTAAGAGACTGATTGGCTT 55

RESULT 4
AA431534 86 bp mRNA EST 22-MAY-1997
LOCUS AA431534/c
DEFINITION zw76h06.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:782171
ACCESSION AA431534
VERSION AA431534.1 GI:2115242
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 86)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT On May 8, 1995 this sequence version replaced gi:800243.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 76.
Location/Qualifiers
1..86
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="782171"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5'

```



```
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE     1 (bases 1 to 82)
AUTHORS       Haller,L., Clark,N., Dubnue,T., Elliston,K., Hawkins,M.,
              Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
              Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
              Trevisakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
              Wilson,R.
TITLE         The Mashu-Merk EST project
JOURNAL       Unpublished (1995)
COMMENT       On May 9, 1995 this sequence version replaced gi:803014.
              Contact: Wilson RK
              Washington University School of Medicine
              444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              Insert Size: 1043
              High quality sequence starts: 1
              Source: IMAGE Consortium, LNLN
              This clone is available royalty-free through LNLN ; contact the
              IMAGE Consortium (info@image.lnl.gov) for further information.
              Trace considered overall poor quality
              Possible reversed clone: similarity on wrong strand
              Insert length: 1043   Std Error: 0.00
              Seq primer: M13RP1
              High quality sequence stop: 1.
FEATURES
SOURCE
    location/Qualifiers
        1..82
        /organism="Homo sapiens"
        /db_xref="GDB:3786918"
        /db_xref="taxon:9606"
        /clone="IMAGE:233741"
        /clone_1lb="Soares fetal liver spleen INFLS"
        /sex="male"
        /dev_stage="20 week-post conception fetus"
        /lab_host="DH10B (ampicillin resistant)"
        /note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACGTGAAGATTAATTAAAGACTCTTTTCTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT
    16 a          19 c          17 g          26 t          4 others
ORIGIN
Query Match      40.0%; Score 16; DB 24; Length 82;
Best Local Similarity 62.9%; Pred.No. 9.4e+03;
Matches 22; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
OY 2 catgtcgcacgtatgagctacaagcgctgatgtt 36
    ||| | | | | | | | | | | | | | | | | |
Db 49 CAGNTAGAAGATGATGACGTGTTTGTCNGTCAYGTG 15
RESULT 8
AA212741/c
LOCUS AA212741 82 bp mRNA EST 18-FEB-1997
DEFINITION mw83c10.r1 Soares mouse N1M Mus musculus cDNA clone IMAGE:677298 5'
similar to SW:YE04.YEAST P32642 HYPOTHEICAL 27.5 KD PROTEIN IN
RA03-BWH1 INTERGENIC REGION.; mRNA sequence.
ACCESSION AA212741 GI:1811370
VERSION AA212741.1
KEYWORDS house mouse.
```

ORGANISM	Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 82)
AUTHORS	Marrá,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Warston,R.
TITLE	The Mashu-HHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	On May 5, 1995 this sequence version replaced gi:797664. Contact: Marrá M/Mouse EST Project Mashu-HHMI Mouse EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mousest@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MG:417002
FEATURES	Trace considered overall poor quality Possible reversed clone: similarity on wrong strand Seq primer: -28ml3 rev2 ET from Amersham High quality sequence stop: 1. location/Qualifiers 1..82
SOURCE	/organism="Mus musculus" /db_xref="taxon:10090" /clone_image=677298" /clone_1ib="Soares mouse NML" /ssuse_type="Liver" /lab_host="DH10B" /note="vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - o1igo(dt) primer [5', TGTTACCAATCTGAAGTGGGAGCGCCGCACAACTTTTTTTTTTTTTTTT 3']; TGTTCACAACTGTAAGTGGGAGCGCCGCACAACTTTTTTTTTTTTTTTT 3'); (Pharmacía), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library constructed and normalized by Bento Soares and M.Fallma Bonaído."
BASE COUNT	18 a 19 c 22 g 23 t
ORIGIN	
Query Match	40.0%; Score 16; DB 30; Length 82;
Best Local Similarity	79.2%; Pred. No. 9.4e+03;
Matches	19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
CY	13 agatgagctacaagcgctgatgttgc 36
DB	24 AAATGACCACACAAGAGTGACTGTG 1
RESUBJ.N	9
LOCUS	H49949/c
DEFINITION	H49949 58 bp mRNA EST 18-SEP-1995
VERSION	y025y03.r1 Soares adult brain NZB5HB55Y Homo sapiens cDNA clone IMAGE:178996 5' similar to SP:SY61_DISOM P24505 SYNAPTOVAGMIN A ;,
KEYWORDS	mRNA sequence.
ACCESSION	H49949 H49949.1 GI:989790
ORGANISM	human.
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 58) Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Holtzman,M., Kucaba,T., Le,M., Lennon,G., Marrá,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,

TITLE Treveskisk, E., Waterston, R., Williamson, A., Wohldmann, P. and
JOURNAL Wilson, R.
COMMENT The Washd-Merck EST Project
Unpublished (1995)
On Apr 14, 1993 this sequence version replaced g1:785348.

FEATURES	Location/Qualifiers
source	1. .58

BASE COUNT	9 a	19 c	17 g	12 t	1 others
ORIGIN					

Query Match	39.5%;	Score 15.8;	DB 23;	Length 58;
Best Local Similarity	71.4%;	Pred. No. 1e+04;		
Matches 20;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0

RESULT	10
A1767746/c	
LOCUS	A1767746 85 bp mRNA
DEFINITION	wh39-007.x1 NC1.CGAP.Kid1 Homo sapiens cDNA clone IMAGE:2383116 3'
	similar to TR:O14099 O14099 HYPOTHETICAL 49.9 KD ZINC FINGER
	PROTEIN C2F3.16 IN CHROMOSOME I., mRNA sequence.
ACCESSION	A1767746
VERSION	G1:5234255
KEYWORDS	EST.
SOURCE	human.

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 85)
REFERENCE	NCI-CCAF http://www.ncbi.nlm.nih.gov/ncicgap .
AUTHORS	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	On May 9, 1996 this sequence version replaced g1.1133104.

FEATURES	Location/Qualifiers
source	1. .85

BASE COUNT	18 a	20 c	17 g	30 t
ORIGIN				

Query Match	39.5%;	Score 15.8;	DB 59;	Length 85;
Best Local Similarity	74.1%;	Pred. No. 1.1e+04;		
Matches 20; Conservative	0;	Mismatches 7;	Indels 0;	Gaps 0

QY	13	agatgagctacaagcgctgtatgga	39
	1		
	1		
	1		
	1		
Db	51	ACATGAGCAACACACGGGATGTGTGA	25
RESULT	11		
LOCUS	R84750	96 bp	MRNA
DEFINITION	Y228a09.r1 Soares fetal liver spleen	INFLU	EST
	IMAGE:274864	5' similar to gb:L12136	HOMO sapiens CDNA C
	(HUMAN); mRNA sequence.	DEOXYCYTIDYLATE	DEAMINASEB

SOURCE ORGANISM	REFERENCE AUTHORS
human.	
<i>Homo sapiens</i>	
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia	
Eutheria: Primates: Catarrhini: Hominoidea: Homo.	
1 (bases 1 to 96)	
Hillier, L., Clark, N., Dubouche, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucabka, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasaks, E., Waterston, R., Williamson, A., Wohlmann, P., and	

TITLE
JOURNAL

The Mashu-Merck EST project
Unpublished (1995)
On Sep 21, 1992 this sequence version replaced gi:276155.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 795

HIGH QUALITY SEQUENCE STARTS: 1
HIGH QUALITY SEQUENCE STOPS: 1
SOURCE: IMAGE Consortium, LNLN
This clone is available royalty-free through LNLN; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
TRACE CONSIDERED OVERALL POOR QUALITY
INSERT LENGTH: 795 STD ERROR: 0.00
SEQ PRIMER: M13RP1
High quality sequence stop: 1.

FEATURES
source location/Qualifiers

1..96
/organism="Homo sapiens"
/db_xref="GDB:3798934"
/db_xref="taxon:9606"
/clone="IMAGE:274864"
/clone_1lb="Soares fetal liver spleen INPLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pTV73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
left strand cDNA was primed with a Pac I - oligo(dT) primer
[5' ACCCGAAGATTAATTAAAGACTCTTTTCTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pTV73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT
ORIGIN

24 a 26 c 21 g 19 t 6 others

Query Match
Best Local Similarity 71.4%; Pred. NO. 1.2e+04;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 gcacgtcgacgtagatgcctacaaggc 28
||||| ||||| | ||||| |||||
DB 41 GCATNCACAGCAGTCATGTCAAGGC 68

RESULT 12
LOCUS w77250/c
DEFINITION me65e02.r1 Soares mouse embryo NDM63.5 14.5 Mus musculus cDNA
clone IMAGE:000442 5' similar to PIR-JC4163 JC4163 DNA-binding
protein 555 - rat ;, mRNA sequence.
W77250 W77250
W77250.1 GI:1387275
EST 20-JUN-1996
mRNA Mus musculus cDNA
w77250 100 bp
EST 20-JUN-1996
mus musculus
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 100)
Marras,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Thelshing,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)

COMMENT

On Apr 14, 1993 this sequence version replaced g1:785288.
Contact: Marra M/Mouse EST Project
WashU-HMHT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LINT ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:244210

FEATURES

source

1..100
location/Qualifiers

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:400442"
/clone_id="Soares mouse embryo NDMEJ3.5 14.5"
/sex="unknown"
/tissue-type="embryo"
/dev-stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/note="vector: pUT7D-Pac (Pharmacia) with a modified
polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAGTAGTGCGGCCGGCCGGAATTGTTCCTTTTTTTTTTTTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2] ; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pUT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT

29 a 28 c 21 g 22 t

ORIGIN

Query Match 39.5%; Score 15.8; DB 26; Length 100;
Best Local Similarity 65.7%; Pred. NO. 1.2e+04;

Matches 23; Conservative 0; Mismatches 12; Indels 0; Gaps 0.

Dy 4 tgtagcgtagattgagctacaaggcgtagttggg 38
| | | | | | | | | | | | | | |
Db 95 TCTTGACGTAGCTGATCTCAAGAAGGCCCTTGC GG 61

RESULT 13

A1360727 71 bp mRNA EST 15-FEB-1999
LOCUS qx38d03.x1 NCI_CGAP.Brn23 Homo sapiens cDNA clone IMAGE:2010533 3'
DEFINITION similar to gb:L07807.DYNAMIN-1 (HUMAN);, mRNA sequence.
ACCESSION A1360727
VERSION A1360727.1 GI:4112348
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 71)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGP), Tumor Gene Index
Unpublished (1998)
On May 8, 1995 this sequence version replaced gi:800368.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov

JOURNAL COMMENT

source

BASE COUNT	7 a	15 c	32 g	16 t
ORIGIN				

Query Match	38.5%;	Score 15.4;	DB 74;	Length 70%;
Best Local Similarity	76.0%;	Pred. No. 1.6e+04;		
Matches 19;	Conservative	0;	Mismatches 6;	Indels

OY 15 atgagctacaagcgtgatltgya 39
 ||| ||| ||||| ||| ||||
Db 46 ATGGGCTTCAAGGCAGGATCAGCGA 22

Search completed: June 4, 2000, 13:53:40
Job time: 20964 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2000, 16:09:48 ; Search time 244.64 Seconds
(without alignments)
21,253 Million cell updates/sec

Title: US-09-164-714-20

Perfect score: 40
Sequence: 1 gcatgtcagctagatgagctacaagcgctattgagat 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 375880

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/5C.COMB.seq:*
4: /cgn2_6/prodata/2/ina/5D.COMB.seq:*
5: /cgn2_6/prodata/2/ina/6.COMB.seq:*
6: /cgn2_6/prodata/2/ina/PCRUS.COMB.seq:*
7: /cgn2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16.6	41.5	25	5	US-08-755-587-157
2	16	40.0	60	1	US-08-180-572-12
3	15.2	38.0	35	1	US-08-347-792-34
4	15.2	38.0	35	2	US-08-431-357-34
5	15.2	38.0	35	6	PCT-US95-15353-34
6	14.8	37.0	88	7	5194596-1
7	14.8	37.0	88	7	5194596-4
8	14.8	37.0	88	7	5219739-1
9	14.8	37.0	88	7	5219739-4
10	14.6	36.5	22	1	US-08-061-889-11
11	14.6	36.5	22	1	US-08-462-611-11
12	14.6	36.5	22	6	PCT-US94-05378-11
13	14.6	36.5	33	2	US-08-538-875-70
14	14.6	36.5	33	4	US-08-600-783-8
15	14.6	36.5	33	4	US-08-600-783-9
16	14.6	36.5	50	3	US-08-867-579-1
17	14.6	36.5	78	3	US-08-659-251-16
18	14.6	36.5	78	6	PCT-US96-11445-16
19	14.4	36.0	45	2	US-08-569-806-15
20	14.4	36.0	45	2	US-08-433-783-24
21	14.4	36.0	46	3	US-08-337-358-24
22	14.4	36.0	46	6	PCT-US93-12144-24
23	14.4	36.0	46	6	PCT-US95-07537A-24
24	14.4	36.0	46	6	PCT-US95-07537-24
25	14.4	36.0	48	6	PCT-US95-02140-5
26	14.2	35.5	65	6	PCT-US95-13376-16
27	14.2	35.5	84	1	US-08-433-126A-191

28	14.2	35.5	84	2	US-08-433-124A-191	Sequence 191, App
29	14.2	35.5	84	6	PCT-US96-06059-191	Sequence 191, App
30	14.2	35.5	99	2	US-08-662-335A-10	Sequence 10, App1
31	14	35.0	33	1	US-07-862-495-7	Sequence 7, App1
32	14	35.0	35	2	US-08-170-290A-36	Sequence 36, App1
33	14	35.0	57	1	US-08-096-027-1	Sequence 1, App1
34	14	35.0	57	2	US-08-461-725-1	Sequence 1, App1
35	14	35.0	57	3	US-08-460-881-1	Sequence 1, App1
36	14	35.0	79	1	US-08-384-708A-217	Sequence 217, App
37	14	35.0	79	3	US-08-470-939-14	Sequence 14, App1
38	14	35.0	79	6	PCT-US96-09452-14	Sequence 14, App1
39	14	35.0	88	7	5194596-6	Patent No. 5194596
40	14	35.0	88	7	5219739-6	Patent No. 5219739
41	13.8	34.5	26	1	US-08-117-362-13	Sequence 13, App1
42	13.8	34.5	26	2	US-08-468-924-13	Sequence 13, App1
43	13.8	34.5	39	2	US-08-468-347-4	Sequence 4, App1
44	13.8	34.5	39	2	US-08-226-264-2	Sequence 2, App1
45	13.8	34.5	39	2	US-08-226-264-11	Sequence 11, App1

ALIGNMENTS

RESULT 1
US-08-755-587-157/C
Sequence 157, Application US/08755587
Patent No. 6045997
GENERAL INFORMATION:
APPLICANT: Futreal, Phillip A
APPLICANT: Wooster, Richard F
APPLICANT: Ashworth, Alan
APPLICANT: Stratton, Michael R
TITLE OF INVENTION: Materials and methods relating to the
TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
NUMBER OF SEQUENCES: 222
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Bell Seltzer Park & Gibson
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
CITY: Raleigh
STATE: NC
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (ERO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9523555.0
FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kenneth D Sibley
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5405-135
INFORMATION FOR SEQ ID NO: 157:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-755-587-157
Query Match 41.5%; Score 16.6; DB 5; Length 25;


```

1 ZIP: 19477
2
3 COMPUTER READABLE FORM:
4
5 MEDIUM TYPE: Floppy disk
6
7 COMPUTER: IBM PC compatible
8
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10 SOFTWARE: PatentIn Release #1.0,
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/431,357
13
14 FILING DATE:
15 CLASSIFICATION: 435
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: US 08/347,792
18 FILING DATE: 28-NOV-1994
19 ATTORNEY/AGENT INFORMATION:
20 NAME: Bak, Mary E.
21 REGISTRATION NUMBER: 31,215
22 REFERENCE/DOCKET NUMBER: WST58USA
23 TELECOMMUNICATION INFORMATION:
24 TELEPHONE: 215-540-9206
25 TELEFAX: 215-540-5818
26 INFORMATION FOR SEQ ID NO: 34:
27 SEQUENCE CHARACTERISTICS:
28 LENGTH: 35 base pairs
29 TYPE: nucleic acid
30 STRANDEDNESS: single
31 TOPOLOGY: linear
32 MOLECULE TYPE: DNA (genomic)
33
34 US-08-431-357-34

```

Query Match	38.0%;	Score 15.2;	DB 2;	Length 35
Best Local Similarity	71.4%;	Pred. No. 3.4e+02;		
Matches	20;	Conservative	0;	Mismatches 8;
				Indels

```

QY      6  tcgacgtagatgagctacaagcgctgat 33
          | | | | | | | | | | | | | |
Db     28  TGGGCATCGATGAGCTCCAAAGGCGCTCAT 1

```

RESULT 5
PCT-US95-15353-34/c
Sequence 34, Application PC/TUS9515353
GENERAL INFORMATION:
APPLICANT: The Wistar Institute of Anatomy
APPLICANT: and Biology
APPLICANT: Halazonetis, Thanos D.
TITLE OF INVENTION: p53 Proteins With Altered
TITLE OF INVENTION: Tetramerization Domains
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., PO Box 455
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15353
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/347,792
FILING DATE: 28-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/431,357
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/456,623

```

1 BILLING DATE: 01-JUN-1995
2
3 ATTORNEY/AGENT INFORMATION:
4
5 NAME: Bak, Mary E.
6
7 REGISTRATION NUMBER: 31,215
8
9 REFERENCE/DOCKET NUMBER: W8T58PCPC
10
11 TELECOMMUNICATION INFORMATION:
12
13 TELEPHONE: 215-540-9206
14
15 TELEFAX: 215-540-5818
16
17 INFORMATION FOR SEQ ID NO: 34:
18
19 SEQUENCE CHARACTERISTICS:
20
21 LENGTH: 35 base pairs
22
23 TYPE: nucleic acid
24
25 STRANDEDNESS: single
26
27 TOPOLOGY: linear
28
29 MOLECULE TYPE: DNA (genomic)
30
31 PCT-US95-13553-34

```

Query Match	38.0%;	Score 15.2;	DB 6;	Length 35;
Best Local Similarity	71.4%;	Pred. No. 3.4e+02;		
Matches 20; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;

Qy	6	tcgacgtagatgagctacaagcgtgat	33
Db	28	TGGGCATCGATGAGCTTCCAAGGCCTCAT	1

RESULT 6
5194596-1/c
Patent No. 5194596
APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES, JOHN
C.; MITCHELL, RICHARD L.
TITLE OF INVENTION: PRODUCTION OF VASCULAR ENDOTHELIAL CELL
GROWTH FACTOR
NUMBER OF SEQUENCES: 32
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/450,883
FILING DATE: 14-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 387,545
FILING DATE: 27-JUL-1989
SEQ ID NO:1:
LENGTH: 88
5194596-1

Query Match	37.0%	Score 14.8	DB 7	Length 88
Best Local Similarity	73.1%	Pred. No. 6	2e+02	
Matches	19	Conservative	0	Mismatches 7
				Indels 0
				Gaps 0
QY	1	gcacgtcgacgtagatgagctacaag	26	
Db	29	GCTGTAGACGTCACATGAACCTTCAAG	4	

```

RESULT 7
5194596-4/c
; Patent No. 5194596
; APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES, JOHN
; C.; MITCHELL, RICHARD L.
; TITLE OF INVENTION: PRODUCTION OF VASCULAR ENDOTHELIAL CELL
; GROWTH FACTOR
; NUMBER OF SEQUENCES: 32
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/450,883
; FILING DATE: 14-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 387,545
; FILING DATE: 27-JUL-1989
; SEQ ID NO:4:
; LENGTH: 88
5194596-4

```

Query Match 37.0%; Score 14.8; DB 7; Length 88;
Best Local Similarity 73.1%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 gcatgtcagcagtagatgagctacaag 26
DB 29 GCTGTAGACGTCCATGAACTTCAAG 4

RESULT 8
5219739-1/c
; Patent No. 5219739
; APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDES,
; JOHN C.; MITCHELL, RICHARD L.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING BVGEF120 AND
; BVGEF 121 AND METHODS FOR THE PRODUCTION OF BOVINE AND HUMAN
; VASCULAR ENDOTHELIAL CELL GROWTH FACTORS, BVGEF120 AND BVGEF121
; NUMBER OF SEQUENCES: 40
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/559,041
; FILING DATE: 27-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 450,883
; FILING DATE: 14-DEC-1989
; APPLICATION NUMBER: 387,545
; FILING DATE: 27-JUL-1989
; SEQ ID NO:1:
; LENGTH: 88
5219739-1

Query Match 37.0%; Score 14.8; DB 7; Length 88;
Best Local Similarity 73.1%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 gcatgtcagcagtagatgagctacaag 26
DB 29 GCTGTAGACGTCCATGAACTTCAAG 4

RESULT 9
5219739-4/c
; Patent No. 5219739
; APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDES,
; JOHN C.; MITCHELL, RICHARD L.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING BVGEF120 AND
; BVGEF 121 AND METHODS FOR THE PRODUCTION OF BOVINE AND HUMAN
; VASCULAR ENDOTHELIAL CELL GROWTH FACTORS, BVGEF120 AND BVGEF121
; NUMBER OF SEQUENCES: 40
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/559,041
; FILING DATE: 27-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 450,883
; FILING DATE: 14-DEC-1989
; APPLICATION NUMBER: 387,545
; FILING DATE: 27-JUL-1989
; SEQ ID NO:4:
; LENGTH: 88
5219739-4

Query Match 37.0%; Score 14.8; DB 7; Length 88;
Best Local Similarity 73.1%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 gcatgtcagcagtagatgagctacaag 26
DB 29 GCTGTAGACGTCCATGAACTTCAAG 4

RESULT 10

US-08-061-889-11/c
; Sequence 11, Application US/08061889
; Patent No. 5654138
; GENERAL INFORMATION:
; APPLICANT: LERMAN, MICHAEL I., LATIF,
; APPLICANT: FARIDA, SZBAR, BERTON, LINEHAN, MARSTON
; TITLE OF INVENTION: VON HIPPEL-LINDAU (VHL)
; TITLE OF INVENTION: DISEASE GENE AND CORRESPONDING CDNA AND
; TITLE OF INVENTION: METHODS FOR DETECTING CARRIERS OF THE VHL
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/061,889
; FILING DATE: 14-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-061-889-11

Query Match 36.5%; Score 14.6; DB 1; Length 22;
Best Local Similarity 81.0%; Pred. No. 5.5e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 18 agctacaagcggtgattggg 38
DB 21 AGCTACAAGCATCACTGTGTG 1

RESULT 11
US-08-462-611-11/c
; Sequence 11, Application US/08462611
; Patent No. 5759790
; GENERAL INFORMATION:
; APPLICANT: LERMAN, MICHAEL I., LATIF,
; APPLICANT: FARIDA, SZBAR, BERTON, LINEHAN, MARSTON
; TITLE OF INVENTION: VON HIPPEL-LINDAU (VHL)
; TITLE OF INVENTION: DISEASE GENE AND CORRESPONDING CDNA AND
; TITLE OF INVENTION: METHODS FOR DETECTING CARRIERS OF THE VHL
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,611
FILING DATE: 05-JUNE-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/061,889
FILING DATE: 14-MAY-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-40780S1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-462-611-11

Query Match 36.5%; Score 14.6; DB 2; Length 22;
Best Local Similarity 81.0%; Pred. No. 5.5e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 agctacaagcgatgattggg 38
|||||
Db 21 AGCTACAAGCATCATCTGTG 1

RESULT 12
PCT-US94-05378-11/C
Sequence 11, Application PC/TUS9405378
GENERAL INFORMATION:
APPLICANT: THE GOVERNMENT OF THE UNITED
STATES OF AMERICA AS REPRESENTED BY THE
APPLICANT: SECRETARY, DEPARTMENT OF HEALTH AND HUMAN
SERVICES
TITLE OF INVENTION: VON HIPPEL-LINDAU (VHL)
TITLE OF INVENTION: DISEASE GENE AND CORRESPONDING CDNA AND
TITLE OF INVENTION: METHODS FOR DETECTING CARRIERS OF THE VHL
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05378
FILING DATE: 12-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/061,889
FILING DATE: 14-MAY-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459

REFERENCE/DOCKET NUMBER: 2026-4078PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US94-05378-11

Query Match 36.5%; Score 14.6; DB 6; Length 22;
Best Local Similarity 81.0%; Pred. No. 5.5e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 agctacaagcgatgattggg 38
|||||
Db 21 AGCTACAAGCATCATCTGTG 1

RESULT 13
US-08-538-875-70/C
Sequence 70, Application US/08538875
Patent No. 5773582
GENERAL INFORMATION:
APPLICANT: Shin, Hang-Cheol
APPLICANT: Shin, Nam-Kyu
APPLICANT: Lee, Inkyung
TITLE OF INVENTION: TUMOR NECROSIS FACTOR MUTAINS
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Shin, Hang-Cheol
STREET: Jukong Gocheung Apt. 1014-806, Haan-dong
CITY: Kwangmyung-shi
STATE: Kyungki-do
COUNTRY: Republic of Korea
ZIP: 423-060
ADDRESSEE: Shin, Nam-Kyu
STREET: #181-404 Sadang-4-dong, Dongjak-ku
CITY: Seoul
STATE:
COUNTRY: Republic of Korea
ZIP: 156-094
ADDRESSEE: Lee, Inkyung
STREET: 11/2, #302-39 Juan-4-dong, Nam-ku
CITY: Incheon
STATE:
COUNTRY: Republic of Korea
ZIP: 402-204
ADDRESSEE: Kang, Sungzong
STREET: #84-4 Daesin-dong, Seodaemun-ku
CITY: Seoul
STATE:
COUNTRY: Republic of Korea
ZIP: 120-160
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5inch 2.0mb storage
COMPUTER: IBM PC/AT
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/538,875
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/193,336
FILING DATE:
APPLICATION NUMBER: KR 93-1751
FILING DATE: 9-FEB-1993

```

? ATTORNEY/AGENT INFORMATION:
? NAME:
? REGISTRATION NUMBER:
? REFERENCE/DOCKET NUMBER:
? TELECOMMUNICATION INFORMATION:
? TELEPHONE:
? TELEFAX:
? INFORMATION FOR SEQ ID NO: 70:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 33 bases
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: primer DNA
?
?
? US-08-538-875-70

```

Query Match 36.5%; Score 14.6; DB 2; Length 33;
 Best Local Similarity 69.0%; Pred.No. 6e+02;
 Matches 20; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 11 gtatgtgagctacaaggcgctgattggga 39
 ||||| ||| | |||||
 db 33 gtatgtgaggaacacagccctctattggga 5

RESULT 14
US-08-600-783-8/c
Sequence 8, Application US/08600783
Patent No. 5962267
GENERAL INFORMATION:
APPLICANT: SHIN, Hang Cheol
APPLICANT: CHANG, Seung Gu
APPLICANT: KIM, Dae Young
APPLICANT: KIM, Chong Suh1
TITLE OF INVENTION: Polysulin Derivative and Process
TITLE OF INVENTION: for Producing Human Insulin
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: SHIN, Hang Cheol
STREET: Saangma-Hanshin Apt. 102-1206,
STREET: #45 Cholsan-dong,
CITY: Kwangmyung-sh1
STATE: Kyungki-do
COUNTRY: Republic of Korea
ZIP: 423-030
ADDRESSEE: CHANG, Seung Gu
STREET: Hyundai Apt. 71-203, Apkujong-dong,
STREET: Kangnam-ku
CITY: Seoul
STATE: Seoul
COUNTRY: Republic of Korea
ZIP: 135-110
ADDRESSEE: KIM, Dae Young
STREET: Sosa Jukong Apt. 108-202, Sosa Bon-dong,
STREET: Sosa-ku
CITY: Bucheon-sh1
STATE: Kyungki-do
COUNTRY: Republic of Korea
ZIP: 422-230
ADDRESSEE: KIM, Chong Suh1
STREET: Garden Heights Apt. 202-801, #100,
STREET: Hwangjeun-dong, Soosung-ku
CITY: Taegu
STATE: Taegu
COUNTRY: Republic of Korea
ZIP: 706-040
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk, 3.5 inch, 1.44MB storage
COMPUTER: IBM PC/AT
OPERATING SYSTEM: MS-DOS
SOFTWARE: word Perfect 5.1
CURRENT APPLICATION DATA:

```

1 APPLICATION NUMBER: US/08/600,783
2
3 FILING DATE:
4
5 CLASSIFICATION: 514
6
7 PRIOR APPLICATION DATA:
8
9 APPLICATION NUMBER: KR 95-2751
10
11 FILING DATE: 15-FEB-1995
12
13 ATTORNEY/AGENT INFORMATION:
14
15 NAME: Shaban Islam
16
17 REGISTRATION NUMBER: 32,507
18
19 REFERENCE/DOCKET NUMBER:
20
21 TELECOMMUNICATION INFORMATION:
22
23 TELEPHONE: (212) 278-1000
24
25 TELEFAX: (212) 953-7249
26
27 INFORMATION FOR SEQ. ID NO: 8:
28
29 SEQUENCE CHARACTERISTICS:
30
31 LENGTH: 33 bases
32
33 TYPE: nucleic acid
34
35 STRANDEDNESS: single
36
37 TOPOLOGY: linear
38
39 MOLECULE TYPE: primer DNA
40
41 ANTI-SENSE: no
42
43 US-08-600-783-8

```

```

Query Match          36.5%; Score 14.6; DB 4; Length 33;
Best Local Similarity 69.0%; Pred. No. 6e+02;
Matches 20; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY      11  gtatagtgcctacaagcgctgattggga 39
          |||||  |||  |  |||||  |
Db       33  GTAGATGAGGAACAGCCCTCTATTGGCA 5

RESULT 15
US-08-600-783-9
; Sequence 9, Application US/08600783
; Patent No. 5962267
; GENERAL INFORMATION:
; APPLICANT: SHIN, Hang Cheol
; APPLICANT: CHANG, Seung Gu
; APPLICANT: KIM, Dae Young
; APPLICANT: KIM, Chong Suh1
; TITLE OF INVENTION: Prolinsulin Derivative and Process
; TITLE OF INVENTION: for Producing Human Insulin
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHIN, Hang Cheol
; STREET: Ssangma-Hanshin Apt. 102-1206,
; STREET: #245 ChoiSan-dong
; CITY: Kwangmyung-shi
; STATE: Kyungki-do
; COUNTRY: Republic of Korea
; ZIP: 423-030
; ADDRESSEE: CHANG, Seung Gu
; STREET: Hyundai Apt. 71-203, ApKujong-dong,
; STREET: Kangnam-ku
; CITY: Seoul
; STATE: Seoul
; COUNTRY: Republic of Korea
; ZIP: 135-110
; ADDRESSEE: KIM, Dae Young
; STREET: Sosa Jukong Apt. 108-202, Sosa Bon-dong,
; STREET: Sosa-ku
; CITY: Bucheon-shi
; STATE: Kyungki-do
; COUNTRY: Republic of Korea
; ZIP: 422-230
; ADDRESSEE: KIM, Chong Suh1
; STREET: Garden Heights Apt. 202-801, #100,
; STREET: Hwangkeum-dong, Soosung-ku
; CITY: Taeju
; STATE: Taeju
; COUNTRY: Republic of Korea

```


ZIP: 706-040
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk, 3.5 inch, 1.44MB storage
COMPUTER: IBM PC/AT
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,783
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: KR 95-2751
FILING DATE: 15-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Shahar Islam
REGISTRATION NUMBER: 32,507
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-1000
TELEFAX: (212) 953-7249
INFORMATION FOR SEQ. ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: primer DNA
ANTI-SENSE: yes
US-08-600-783-9

Query Match 36.5%; Score 14.6; DB 4; Length 33;
Best Local Similarity 69.0%; Pred. No. 6e+02;
Matches 20; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 11 gtagatagctacaagcgctcattggga 39
||||||| ||| ||| |||||
Db 1 GTAGATGAGCAGACAGCCCTCTATTGGCA 29

Search completed: June 4, 2000, 16:09:50
Job time: 28069 sec

THIS PAGE BLANK (USPTO)